GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 10:27:46 ; Search time 1299 Seconds
(without alignments)
11496.806 Million cell updates/sec

Title: US-10-028-248A-210
Perfect score: 3260
Sequence: 1 ggcacgagcaggagcttccc.....aaaaaaaaaaaaaaaaaaa 3260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3259	100.0	3260	16	US-10-028-248A-210	Sequence 210, App
2	3259	100.0	3260	16	US-10-107-782-210	Sequence 210, App
3	3232.6	99.2	4576	9	US-09-842-930A-24	Sequence 24, Appl
4	3232.6	99.2	4642	15	US-10-133-172-3	Sequence 3, Appli
5	2999.8	92.0	8444	16	US-10-028-248A-1	Sequence 1, Appli
6	2999.8	92.0	8444	16	US-10-107-782-1	Sequence 1, Appli
7	2995.6	91.9	8495	16	US-10-028-248A-3	Sequence 3, Appli
8	2995.6	91.9	8495	16	US-10-107-782-3	Sequence 3, Appli
9	2853.2	87.5	4962	15	US-10-133-172-19	Sequence 19, Appl
10	1936.8	59.4	4706	9	US-09-842-930A-1	Sequence 1, Appli
11	1936.8	59.4	4706	15	US-10-133-172-1	Sequence 1, Appli
12	708	21.7	718	15	US-10-023-896-38	Sequence 38, Appl
13	537	16.5	722	9	US-09-822-830A-559	Sequence 559, App
14	424.6	13.0	3681	16	US-10-264-237-1320	Sequence 1320, Ap
15	408.8	12.5	448	10	US-09-918-995-30654	Sequence 30654, A
16	391.2	12.0	2863	16	US-10-104-047-610	Sequence 610, App
17	285	8.7	2160	16	US-10-104-047-1824	Sequence 1824, Ap
18	259.6	8.0	2753	10	US-09-774-639-12	Sequence 12, Appl
19	259.6	8.0	2753	10	US-09-969-730-13	Sequence 13, Appl
20	259.6	8.0	2753	17	US-10-621-363-13	Sequence 13, Appl
21	247.4	7.6	395	10	US-09-918-995-16736	Sequence 16736, A
22	187.4	5.7	1482	16	US-10-305-720-230	Sequence 230, App
23	176.2	5.4	293	9	US-09-783-590-5045	Sequence 5045, Ap
24	143	4.4	473	10	US-09-918-995-2410	Sequence 2410, Ap
25	135	4.1	420	10	US-09-918-995-8459	Sequence 8459, Ap
26	125.2	3.8	474	10	US-09-918-995-32777	Sequence 32777, A
27	107.2	3.3	409	9	US-09-833-381-1252	Sequence 1252, Ap
28	82	2.5	466	13	US-10-085-783A-58386	Sequence 58386, A
29	82	2.5	466	16	US-10-242-535A-58386	Sequence 58386, A
30	81.4	2.5	1144	14	US-10-044-090-446	Sequence 446, App
31	81.4	2.5	1144	15	US-10-247-671-124	Sequence 124, App
32	81.4	2.5	1414	9	US-09-799-118-1	Sequence 1, Appli
33	81.4	2.5	1422	16	US-10-295-027-291	Sequence 291, App
34	81.4	2.5	1430	16	US-10-295-027-1020	Sequence 1020, Ap
35	81.4	2.5	1728	14	US-10-044-090-445	Sequence 445, App
36	81.4	2.5	7840	12	US-10-634-574-15	Sequence 15, Appl
C 37	78.2	2.4	351	15	US-10-060-036-880	Sequence 880, App
38	78.2	2.4	1414	16	US-10-133-937-18	Sequence 18, Appl
39	78.2	2.4	1414	16	US-10-295-027-293	Sequence 293, App
40	78.2	2.4	1414	16	US-10-159-563-18	Sequence 18, Appl
41	74	2.3	396	10	US-09-918-995-17767	Sequence 17767, A
42	73.6	2.3	1411	9	US-09-765-231A-44	Sequence 44, Appl
43	60.4	1.9	2627	16	US-10-108-260A-1663	Sequence 1663, Ap
44	60	1.8	60	10	US-09-908-975-13589	Sequence 13589, A
45	58.8	1.8	781	13	US-10-276-774-488	Sequence 488, App

ALIGNMENTS

RESULT 1
US-10-028-248A-210
; Sequence 210, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959

; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 210
; LENGTH: 3260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3171)
; OTHER INFORMATION: Wherein n is A, C, G, or T
US-10-028-248A-210

Query Match 100.0%; Score 3259; DB 16; Length 3260;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGG 60

Qy         61 AGCATTTCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCATTTCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTG 120

Qy        121 CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGG 180

Qy        181 TTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCT 240

Qy        241 CAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGT 300

Qy        301 ATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTC 360

Qy        361 ATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTG 420

Qy        421 GAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCA 480

Qy        481 ACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCA 540

Qy        541 CTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600

Qy        601 TGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAG 660

Qy        661 ATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTT 720

Qy        721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCC 780

Qy        781 AAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTG 840

Qy        841 ACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG 900

Qy        901 CCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAA 960

Qy        961 AGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCC 1020

Qy       1021 GGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAG 1080

Qy       1081 ACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTG 1140

Qy       1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG 1200

Qy       1201 AGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACG 1260

Qy       1261 GACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAG 1320

Qy       1321 GCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTC 1380

Qy       1381 ATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAA 1440

Qy       1441 TCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCA 1500

Qy       1501 GATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACG 1560

Qy       1561 GGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGC 1620

Qy       1621 ACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATG 1680

Qy       1681 TCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACA 1740

Qy       1741 ATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTG 1800

Qy       1801 GGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860
```

```
Db       1801 GGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860
Qy       1861 AGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGA 1920
Db       1861 AGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGA 1920
Qy       1921 AGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCA 1980
Db       1921 AGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCA 1980
Qy       1981 CAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTA 2040
Db       1981 CAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTA 2040
Qy       2041 GACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCA 2100
Db       2041 GACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCA 2100
Qy       2101 CCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCC 2160
Db       2101 CCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCC 2160
Qy       2161 TGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCT 2220
Db       2161 TGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCT 2220
Qy       2221 CAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCT 2280
Db       2221 CAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCT 2280
Qy       2281 TTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGC 2340
Db       2281 TTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGC 2340
Qy       2341 ACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGC 2400
Db       2341 ACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGC 2400
Qy       2401 AAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGG 2460
Db       2401 AAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGG 2460
Qy       2461 AGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGA 2520
Db       2461 AGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGA 2520
Qy       2521 TCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACA 2580
Db       2521 TCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACA 2580
Qy       2581 CTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGG 2640
Db       2581 CTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGG 2640
Qy       2641 CTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGG 2700
Db       2641 CTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGG 2700
Qy       2701 AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCT 2760
Db       2701 AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCT 2760
Qy       2761 TGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTG 2820
Db       2761 TGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTG 2820
Qy       2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880
Db       2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880
Qy       2881 TGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCC 2940
Db       2881 TGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCC 2940
Qy       2941 TTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATC 3000
Db       2941 TTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATC 3000
Qy       3001 TGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCC 3060
Db       3001 TGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCC 3060
Qy       3061 TCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCC 3120
Db       3061 TCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCC 3120
Qy       3121 TACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTAT 3180
Db       3121 TACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTAT 3180
Qy       3181 CCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240
Db       3181 CCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240
Qy       3241 AAAAAAAAAAAAAAAAAAAA 3260
Db       3241 AAAAAAAAAAAAAAAAAAAA 3260
```

RESULT 2
US-10-107-782-210
; Sequence 210, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266

```
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 60/279,344
;  PRIOR FILING DATE: 2001-03-28
;  NUMBER OF SEQ ID NOS: 215
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 210
;   LENGTH: 3260
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (171)..(2861)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (3171)..(3171)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
US-10-107-782-210

  Query Match           100.0%;  Score 3259;  DB 16;  Length 3260;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3260;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGG 60

Qy         61 AGCATTTCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCATTTCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTG 120

Qy        121 CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGG 180

Qy        181 TTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCT 240

Qy        241 CAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGT 300

Qy        301 ATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTC 360

Qy        361 ATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTG 420

Qy        421 GAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCA 480

Qy        481 ACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCA 540

Qy        541 CTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600

Qy        601 TGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAG 660

Qy        661 ATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTT 720

Qy        721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCC 780

Qy        781 AAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTG 840

Qy        841 ACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG 900

Qy        901 CCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAA 960

Qy        961 AGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCC 1020

Qy       1021 GGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAG 1080

Qy       1081 ACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTG 1140

Qy       1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG 1200

Qy       1201 AGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACG 1260

Qy       1261 GACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAG 1320

Qy       1321 GCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTC 1380

Qy       1381 ATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAA 1440

Qy       1441 TCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCA 1500

Qy       1501 GATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACG 1560

Qy       1561 GGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGC 1620

Qy       1621 ACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATG 1680

Qy       1681 TCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACA 1740

Qy       1741 ATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTG 1800

Qy       1801 GGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860
```

```
Qy       1861 AGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGA 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGA 1920

Qy       1921 AGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCA 1980

Qy       1981 CAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTA 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTA 2040

Qy       2041 GACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCA 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCA 2100

Qy       2101 CCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCC 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCC 2160

Qy       2161 TGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCT 2220

Qy       2221 CAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCT 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCT 2280

Qy       2281 TTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGC 2340

Qy       2341 ACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGC 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGC 2400

Qy       2401 AAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGG 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGG 2460

Qy       2461 AGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGA 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 AGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGA 2520

Qy       2521 TCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACA 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACA 2580

Qy       2581 CTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGG 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGG 2640

Qy       2641 CTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGG 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGG 2700

Qy       2701 AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCT 2760

Qy       2761 TGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTG 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTG 2820

Qy       2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880

Qy       2881 TGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCC 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCC 2940

Qy       2941 TTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATC 3000

Qy       3001 TGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCC 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCC 3060

Qy       3061 TCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCC 3120

Qy       3121 TACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTAT 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTAT 3180

Qy       3181 CCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240

Qy       3241 AAAAAAAAAAAAAAAAAAAA 3260
             ||||||||||||||||||||
Db       3241 AAAAAAAAAAAAAAAAAAAA 3260
```

RESULT 3
US-09-842-930A-24
; Sequence 24, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.603
; CURRENT APPLICATION NUMBER: US/09/842,930A
; CURRENT FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 4576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-842-930A-24

Query Match 99.2%; Score 3232.6; DB 9; Length 4576;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3238; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 1387

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 1447

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1448 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 1507

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1508 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 1567

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1568 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 1627

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1628 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 1687
```

```
Qy       367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
Db      1688 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 1747

Qy       427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
Db      1748 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 1807

Qy       487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
Db      1808 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 1867

Qy       547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
Db      1868 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 1927

Qy       607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
Db      1928 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 1987

Qy       667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
Db      1988 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 2047

Qy       727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
Db      2048 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 2107

Qy       787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
Db      2108 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 2167

Qy       847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
Db      2168 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 2227

Qy       907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
Db      2228 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 2287

Qy       967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
Db      2288 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 2347

Qy      1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
Db      2348 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 2407

Qy      1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
Db      2408 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 2467

Qy      1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
Db      2468 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 2527

Qy      1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
Db      2528 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 2587

Qy      1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
Db      2588 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 2647

Qy      1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
Db      2648 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 2707

Qy      1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
Db      2708 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 2767

Qy      1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
Db      2768 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 2827

Qy      1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
Db      2828 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 2887

Qy      1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
Db      2888 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 2947

Qy      1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
Db      2948 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 3007

Qy      1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
Db      3008 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 3067

Qy      1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
Db      3068 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 3127

Qy      1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
Db      3128 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 3187

Qy      1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
Db      3188 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 3247

Qy      1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
Db      3248 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 3307

Qy      1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
Db      3308 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 3367

Qy      2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
Db      3368 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 3427

Qy      2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
Db      3428 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 3487

Qy      2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
Db      3488 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 3547

Qy      2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
Db      3548 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 3607

Qy      2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
Db      3608 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 3667

Qy      2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
Db      3668 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 3727

Qy      2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
Db      3728 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 3787

Qy      2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
Db      3788 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 3847

Qy      2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
```

```
Db       3848 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 3907

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646

Db       3908 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 3967

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706

Db       3968 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 4027

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766

Db       4028 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 4087

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826

Db       4088 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 4147

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886

Db       4148 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 4207

Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946

Db       4208 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 4267

Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006

Db       4268 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 4327

Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066

Db       4328 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 4387

Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126

Db       4388 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 4447

Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186

Db       4448 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 4507

Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246

Db       4508 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 4567

Qy       3247 AAAAAAAA 3254

Db       4568 AGTCAACA 4575
```

```
RESULT 4
US-10-133-172-3
; Sequence 3, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
;  APPLICANT: WEIGEL, PAUL H
;  APPLICANT:  WEIGEL, JANET A
;  TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
;  FILE REFERENCE: 5864.014
;  CURRENT APPLICATION NUMBER: US/10/133,172
;  CURRENT FILING DATE:  2002-04-25
;  PRIOR APPLICATION NUMBER: 60/286,468
;  PRIOR FILING DATE: 2001-04-25
;  PRIOR APPLICATION NUMBER: 09/842,930
;  PRIOR FILING DATE: 2001-04-25
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 4642
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

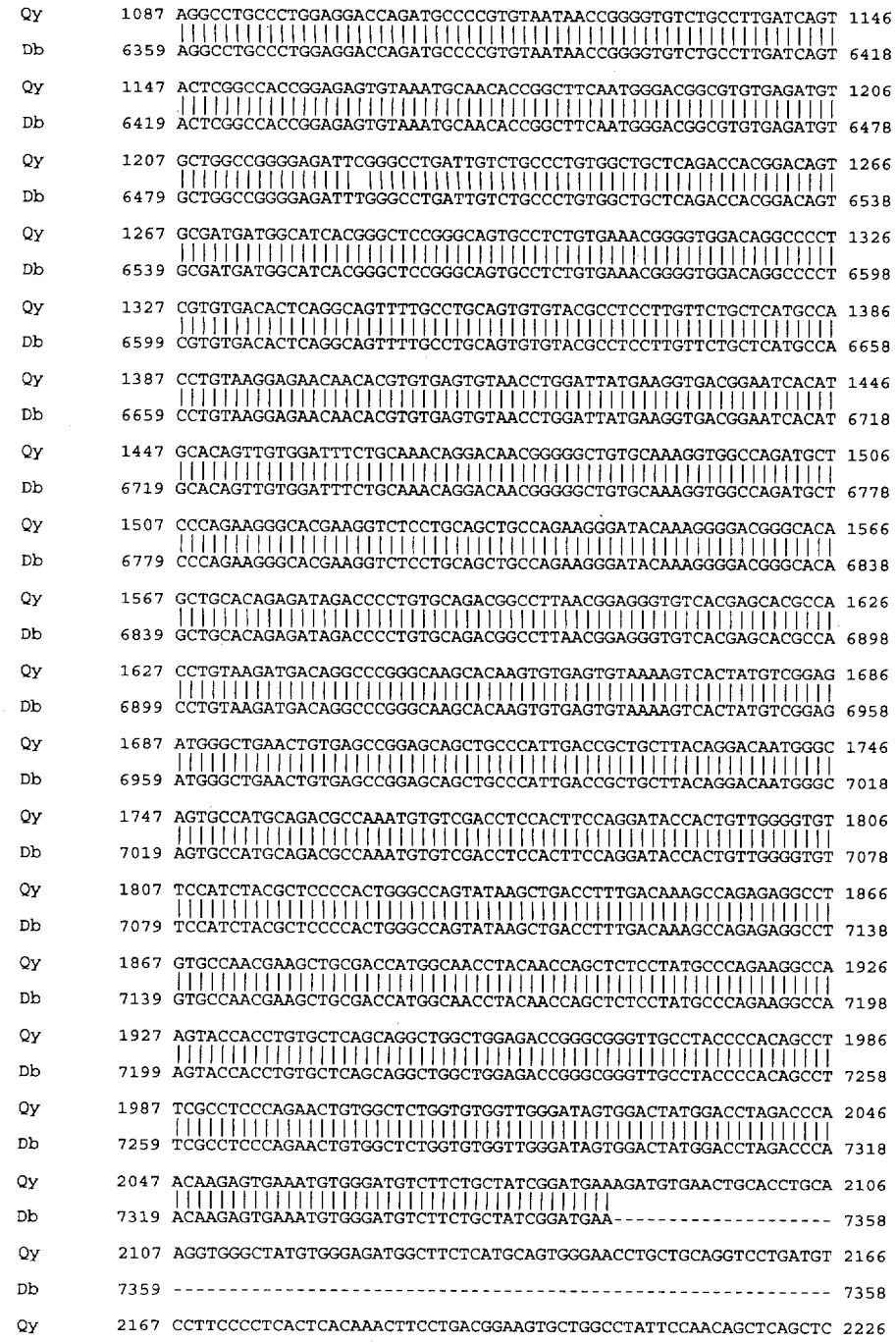
US-10-133-172-3

Query Match 99.2%; Score 3232.6; DB 15; Length 4642;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3238; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66

Db       1394 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 1453

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126

Db       1454 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 1513

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186

Db       1514 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 1573

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246

Db       1574 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 1633

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306

Db       1634 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 1693

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366

Db       1694 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 1753

Qy        367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426

Db       1754 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 1813

Qy        427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486

Db       1814 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 1873

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546

Db       1874 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 1933

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606

Db       1934 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 1993

Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666

Db       1994 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 2053

Qy        667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726

Db       2054 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 2113

Qy        727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786

Db       2114 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 2173

Qy        787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846

Db       2174 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 2233

Qy        847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906

Db       2234 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 2293

Qy        907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966

Db       2294 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 2353

Qy        967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026

Db       2354 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 2413
```

```
Qy    1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2414 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 2473

Qy    1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2474 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 2533

Qy    1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2534 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 2593

Qy    1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    2594 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 2653

Qy    1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2654 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 2713

Qy    1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2714 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 2773

Qy    1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2774 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 2833

Qy    1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2834 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 2893

Qy    1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2894 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 2953

Qy    1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2954 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 3013

Qy    1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3014 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 3073

Qy    1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3074 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 3133

Qy    1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3134 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 3193

Qy    1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3194 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 3253

Qy    1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3254 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 3313

Qy    1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3314 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 3373

Qy    1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3374 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 3433

Qy    2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3434 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 3493

Qy    2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3494 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 3553

Qy    2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3554 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 3613

Qy    2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3614 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 3673

Qy    2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3674 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 3733

Qy    2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3734 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 3793

Qy    2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3794 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 3853

Qy    2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3854 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 3913

Qy    2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3914 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 3973

Qy    2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3974 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 4033

Qy    2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4034 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 4093

Qy    2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4094 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 4153

Qy    2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4154 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 4213

Qy    2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4214 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 4273

Qy    2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4274 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 4333

Qy    2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4334 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 4393

Qy    3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4394 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 4453

Qy    3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4454 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 4513

Qy    3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186
           ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    4514 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 4573

Qy    3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246
```

```
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db      4574 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 4633

Qy      3247 AAAAAAAA 3254
             |   || |
Db      4634 AGTCAACA 4641


RESULT 5
US-10-028-248A-1
; Sequence 1, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
;  APPLICANT: Shimkets, Richard
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Gangolli, Esha
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Taupier Jr, Raymond J
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Colman, Steven
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Si, Jingsheng
;  APPLICANT:  Edinger, Shlomit
;  APPLICANT:  Stone, David
;  APPLICANT:  Sciore, Paul
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Rothenberg, Mark
;  TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 21402-222
;  CURRENT APPLICATION NUMBER: US/10/028,248A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/256619
;  PRIOR FILING DATE: 2000-12-19
;  PRIOR APPLICATION NUMBER: 60/262959
;  PRIOR FILING DATE: 2001-01-19
;  PRIOR APPLICATION NUMBER: 60/272408
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/285189
;  PRIOR FILING DATE: 2001-04-20
;  PRIOR APPLICATION NUMBER: 60/308039
;  PRIOR FILING DATE: 2001-07-26
;  PRIOR APPLICATION NUMBER: 60/311266
;  PRIOR FILING DATE: 2001-08-09
;  NUMBER OF SEQ ID NOS: 211
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 8444
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (163)..(168)
;   OTHER INFORMATION: wherein n is A, C, G, or T
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1279)..(1284)
;   OTHER INFORMATION: wherein n is A, C, G, or T
US-10-028-248A-1

  Query Match             92.0%;  Score 2999.8;  DB 16;  Length 8444;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches 3128;  Conservative    0;  Mismatches   18;  Indels  108;  Gaps    1;

Qy         7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5279 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5338

Qy        67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5339 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 5398

Qy       127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5399 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 5458

Qy       187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5459 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 5518

Qy       247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5519 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 5578

Qy       307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5579 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 5638

Qy       367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5639 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 5698

Qy       427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5699 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 5758

Qy       487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5759 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 5818

Qy       547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5819 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 5878

Qy       607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5879 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 5938

Qy       667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5939 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 5998

Qy       727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5999 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 6058

Qy       787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6059 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 6118

Qy       847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6119 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 6178

Qy       907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6179 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 6238

Qy       967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6239 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 6298

Qy      1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6299 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 6358
```


Qy 1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
Db 6359 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 6418
Qy 1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
Db 6419 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 6478
Qy 1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
Db 6479 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 6538
Qy 1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
Db 6539 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 6598
Qy 1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
Db 6599 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 6658
Qy 1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
Db 6659 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 6718
Qy 1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
Db 6719 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 6778
Qy 1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
Db 6779 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 6838
Qy 1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
Db 6839 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 6898
Qy 1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
Db 6899 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 6958
Qy 1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
Db 6959 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 7018
Qy 1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
Db 7019 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 7078
Qy 1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
Db 7079 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 7138
Qy 1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
Db 7139 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 7198
Qy 1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
Db 7199 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 7258
Qy 1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
Db 7259 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 7318
Qy 2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
Db 7319 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAA---------- 7358
Qy 2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
Db 7359 ---------- 7358
Qy 2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
Db 7359 ----------GGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 7390
Qy 2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
Db 7391 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 7450
Qy 2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
Db 7451 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 7510
Qy 2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
Db 7511 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 7570
Qy 2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
Db 7571 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGGTACAAAGTA 7630
Qy 2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
Db 7631 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 7690
Qy 2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
Db 7691 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 7750
Qy 2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
Db 7751 TGGGAGCAGGGATCTTCTTTTGCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 7810
Qy 2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
Db 7811 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTACCAGCATTTTGAGTCGGAAGAGG 7870
Qy 2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
Db 7871 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 7930
Qy 2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
Db 7931 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 7990
Qy 2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
Db 7991 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 8050
Qy 2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
Db 8051 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 8110
Qy 2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
Db 8111 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 8170
Qy 3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
Db 8171 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 8230
Qy 3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126
Db 8231 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 8290
Qy 3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186
Db 8291 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 8350
Qy 3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246
Db 8351 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 8410
Qy 3247 AAAAAAAAAAAAA 3260

```
Db       8411 AGTCAAAAAAAAAA 8424


RESULT 6
US-10-107-782-1
; Sequence 1, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
;  APPLICANT: Boldog, Ferenc,
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Colman, Steve,
;  APPLICANT:  Edinger, Shlomit,
;  APPLICANT:  Gangolli, Esha,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Liu, Xiaohong,
;  APPLICANT:  Malyankar, Uriel,
;  APPLICANT:  Miller, Charles,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rothenberg, Mark,
;  APPLICANT:  Sciore, Paul,
;  APPLICANT:  Shenoy, Suresh,
;  APPLICANT:  Shimkets, Richard,
;  APPLICANT:  Si, Jingsheng,
;  APPLICANT:  Smithson, Glennda,
;  APPLICANT:  Spytek, Kimberly,
;  APPLICANT:  Stone, David,
;  APPLICANT:  Taupier, Raymond, jr.,
;  APPLICANT:  Tchernev, Velizar,
;  APPLICANT:  Vernet, Corine,
;  APPLICANT:  Zerhusen, Brian
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
;  FILE REFERENCE: 21402-222CIP
;  CURRENT APPLICATION NUMBER: US/10/107,782
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: 10/028,248
;  PRIOR FILING DATE: 2001-12-19
;  PRIOR APPLICATION NUMBER: 60/256,619
;  PRIOR FILING DATE: 2000-12-19
;  PRIOR APPLICATION NUMBER: 60/262,959
;  PRIOR FILING DATE: 2001-01-19
;  PRIOR APPLICATION NUMBER: 60/272,408
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/285,189
;  PRIOR FILING DATE: 2001-04-20
;  PRIOR APPLICATION NUMBER: 60/308,039
;  PRIOR FILING DATE: 2001-07-26
;  PRIOR APPLICATION NUMBER: 60/311,266
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 60/279,344
;  PRIOR FILING DATE: 2001-03-28
;  NUMBER OF SEQ ID NOS: 215
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1
;   LENGTH: 8444
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(8025)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (163)..(163)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (164)..(164)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (165)..(165)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (166)..(166)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (167)..(167)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (168)..(168)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1279)..(1279)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1280)..(1280)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1281)..(1281)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1282)..(1282)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1283)..(1283)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1284)..(1284)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (163)..(165)
;   OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (166)..(168)
;   OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1279)..(1281)
;   OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1282)..(1284)
;   OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-10-107-782-1

  Query Match             92.0%;  Score 2999.8;  DB 16;  Length 8444;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches 3128;  Conservative    0;  Mismatches   18;  Indels  108;  Gaps    1;

Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5279 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5338

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5339 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 5398

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5399 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 5458

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5459 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 5518
```

```
Qy       247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5519 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 5578

Qy       307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5579 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 5638

Qy       367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5639 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 5698

Qy       427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5699 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 5758

Qy       487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5759 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 5818

Qy       547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5819 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 5878

Qy       607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5879 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 5938

Qy       667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5939 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 5998

Qy       727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5999 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 6058

Qy       787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6059 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 6118

Qy       847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6119 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 6178

Qy       907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6179 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 6238

Qy       967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6239 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 6298

Qy      1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6299 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 6358

Qy      1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6359 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 6418

Qy      1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6419 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 6478

Qy      1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6479 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 6538

Qy      1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6539 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 6598
```

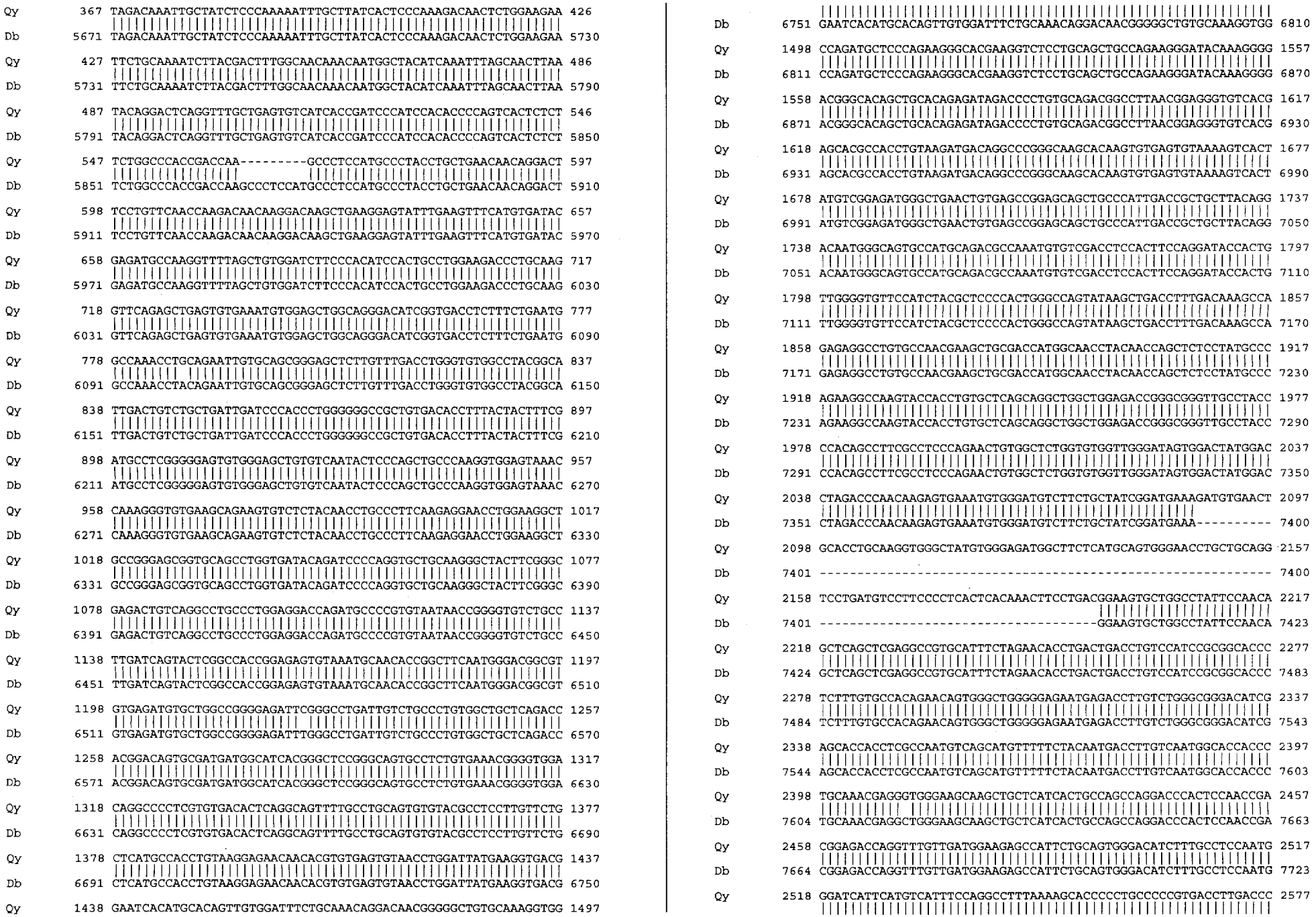
```
Qy      1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6599 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 6658

Qy      1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6659 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 6718

Qy      1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6719 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 6778

Qy      1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6779 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 6838

Qy      1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6839 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 6898

Qy      1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6899 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 6958

Qy      1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6959 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 7018

Qy      1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7019 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 7078

Qy      1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7079 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 7138

Qy      1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7139 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 7198

Qy      1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7199 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 7258

Qy      1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7259 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 7318

Qy      2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7319 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAA---------------------- 7358

Qy      2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7359 ---------------------------------------------------------- 7358

Qy      2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
                                                ||||||||||||||||||||||||||||
Db      7359 ----------------------------GGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 7390

Qy      2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7391 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 7450

Qy      2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7451 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 7510

Qy      2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7511 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 7570

Qy      2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
```

```
Db       7571 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGGTACAAAGTA 7630

Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526

Db       7631 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 7690

Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586

Db       7691 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 7750

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646

Db       7751 TGGGAGCAGGGATCTTCTTTTGCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 7810

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706

Db       7811 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTACCAGCATTTTGAGTCGGAAGAGG 7870

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766

Db       7871 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 7930

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826

Db       7931 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 7990

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886

Db       7991 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 8050

Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946

Db       8051 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 8110

Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006

Db       8111 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 8170

Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066

Db       8171 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 8230

Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126

Db       8231 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 8290

Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186

Db       8291 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 8350

Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246

Db       8351 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 8410

Qy       3247 AAAAAAAAAAAAA 3260

Db       8411 AGTCAAAAAAAAA 8424
```

RESULT 7
US-10-028-248A-3
; Sequence 3, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods'
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 8495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1311)..(1316)
; OTHER INFORMATION: Wherein n is A, C, G, or T
US-10-028-248A-3

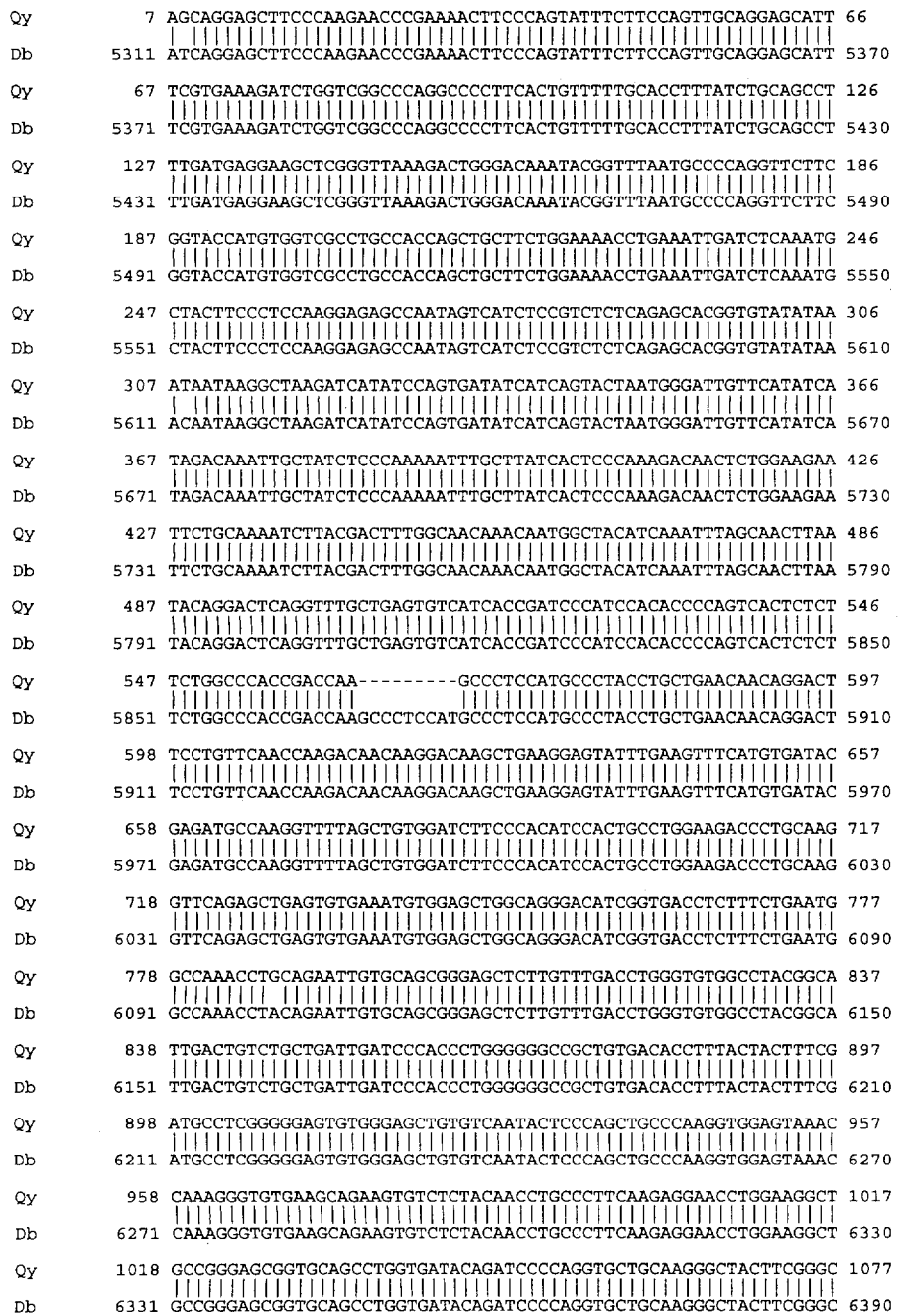
Query Match 91.9%; Score 2995.6; DB 16; Length 8495;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 3137; Conservative 0; Mismatches 10; Indels 116; Gaps 2;

```
Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66

Db       5311 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5370

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126

Db       5371 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 5430

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186

Db       5431 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 5490

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246

Db       5491 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 5550

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306

Db       5551 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 5610

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366

Db       5611 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 5670
```


Qy 367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
Db 5671 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 5730
Qy 427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
Db 5731 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 5790
Qy 487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
Db 5791 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 5850
Qy 547 TCTGGCCCACCGACCAA---------GCCCTCCATGCCCTACCTGCTGAACAACAGGACT 597
Db 5851 TCTGGCCCACCGACCAAGCCCTCCATGCCCTCCATGCCCTACCTGCTGAACAACAGGACT 5910
Qy 598 TCCTGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATAC 657
Db 5911 TCCTGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATAC 5970
Qy 658 GAGATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAG 717
Db 5971 GAGATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAG 6030
Qy 718 GTTCAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATG 777
Db 6031 GTTCAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATG 6090
Qy 778 GCCAAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCA 837
Db 6091 GCCAAACCTACAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCA 6150
Qy 838 TTGACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCG 897
Db 6151 TTGACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCG 6210
Qy 898 ATGCCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAAC 957
Db 6211 ATGCCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAAC 6270
Qy 958 CAAAGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCT 1017
Db 6271 CAAAGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCT 6330
Qy 1018 GCCGGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGC 1077
Db 6331 GCCGGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGC 6390
Qy 1078 GAGACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCC 1137
Db 6391 GAGACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCC 6450
Qy 1138 TTGATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGT 1197
Db 6451 TTGATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGT 6510
Qy 1198 GTGAGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACC 1257
Db 6511 GTGAGATGTGCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACC 6570
Qy 1258 ACGGACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGA 1317
Db 6571 ACGGACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGA 6630
Qy 1318 CAGGCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTG 1377
Db 6631 CAGGCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTG 6690
Qy 1378 CTCATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACG 1437
Db 6691 CTCATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACG 6750
Qy 1438 GAATCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGG 1497
Db 6751 GAATCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGG 6810
Qy 1498 CCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGG 1557
Db 6811 CCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGG 6870
Qy 1558 ACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACG 1617
Db 6871 ACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACG 6930
Qy 1618 AGCACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACT 1677
Db 6931 AGCACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACT 6990
Qy 1678 ATGTCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGG 1737
Db 6991 ATGTCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGG 7050
Qy 1738 ACAATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTG 1797
Db 7051 ACAATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTG 7110
Qy 1798 TTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCA 1857
Db 7111 TTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCA 7170
Qy 1858 GAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCC 1917
Db 7171 GAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCC 7230
Qy 1918 AGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACC 1977
Db 7231 AGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACC 7290
Qy 1978 CCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGAC 2037
Db 7291 CCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGAC 7350
Qy 2038 CTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACT 2097
Db 7351 CTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAA---------- 7400
Qy 2098 GCACCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGG 2157
Db 7401 ---------- 7400
Qy 2158 TCCTGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACA 2217
Db 7401 ----------GGAAGTGCTGGCCTATTCCAACA 7423
Qy 2218 GCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCC 2277
Db 7424 GCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCC 7483
Qy 2278 TCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCG 2337
Db 7484 TCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCG 7543
Qy 2338 AGCACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCC 2397
Db 7544 AGCACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCC 7603
Qy 2398 TGCAAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGA 2457
Db 7604 TGCAAACGAGGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGA 7663
Qy 2458 CGGAGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATG 2517
Db 7664 CGGAGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATG 7723
Qy 2518 GGATCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCC 2577

```
Db       7724 GGATCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCC 7783

Qy       2578 ACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCT 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7784 ACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCT 7843

Qy       2638 TGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGT 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7844 TGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGT 7903

Qy       2698 CGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7904 CGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACC 7963

Qy       2758 CCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACT 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7964 CCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACT 8023

Qy       2818 CTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGG 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8024 CTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGG 8083

Qy       2878 AGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTT 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8084 AGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTT 8143

Qy       2938 GCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 2997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8144 GCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 8203

Qy       2998 ATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCC 3057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8204 ATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCC 8263

Qy       3058 TCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTG 3117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8264 TCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTG 8323

Qy       3118 CCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTG 3177
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       8324 CCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTG 8383

Qy       3178 TATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAA 3237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8384 TATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAA 8443

Qy       3238 CAGAAAAAAAAAAAAAAAAAAA 3260
             |||||| |||   ||||||||||
Db       8444 CAGAAACAAAGTCAAAAAAAAAA 8466
```

```
RESULT 8
US-10-107-782-3
; Sequence 3, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
;  APPLICANT: Boldog, Ferenc,
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Colman, Steve,
;  APPLICANT:  Edinger, Shlomit,
;  APPLICANT:  Gangolli, Esha,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Liu, Xiaohong,
;  APPLICANT:  Malyankar, Uriel,
;  APPLICANT:  Miller, Charles,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rothenberg, Mark,
;  APPLICANT:  Sciore, Paul,
;  APPLICANT:  Shenoy, Suresh,
;  APPLICANT:  Shimkets, Richard,
;  APPLICANT:  Si, Jingsheng,
;  APPLICANT:  Smithson, Glennda,
;  APPLICANT:  Spytek, Kimberly,
;  APPLICANT:  Stone, David,
;  APPLICANT:  Taupier, Raymond, jr.,
;  APPLICANT:  Tchernev, Velizar,
;  APPLICANT:  Vernet, Corine,
;  APPLICANT:  Zerhusen, Brian
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
;  FILE REFERENCE: 21402-222CIP
;  CURRENT APPLICATION NUMBER: US/10/107,782
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: 10/028,248
;  PRIOR FILING DATE: 2001-12-19
;  PRIOR APPLICATION NUMBER: 60/256,619
;  PRIOR FILING DATE: 2000-12-19
;  PRIOR APPLICATION NUMBER: 60/262,959
;  PRIOR FILING DATE: 2001-01-19
;  PRIOR APPLICATION NUMBER: 60/272,408
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/285,189
;  PRIOR FILING DATE: 2001-04-20
;  PRIOR APPLICATION NUMBER: 60/308,039
;  PRIOR FILING DATE: 2001-07-26
;  PRIOR APPLICATION NUMBER: 60/311,266
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 60/279,344
;  PRIOR FILING DATE: 2001-03-28
;  NUMBER OF SEQ ID NOS: 215
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 3
;   LENGTH: 8495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(7460)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1311)..(1311)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1312)..(1312)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1313)..(1313)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1314)..(1314)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1315)..(1315)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1316)..(1316)
;   OTHER INFORMATION: Wherein n may be a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1311)..(1313)
;   OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1314)..(1316)
;   OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-10-107-782-3

  Query Match           91.9%;  Score 2995.6;  DB 16;  Length 8495;
  Best Local Similarity  96.1%;  Pred. No. 0;
  Matches 3137;  Conservative    0;  Mismatches   10;  Indels  116;  Gaps    2;
```


Qy 7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
Db 5311 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5370
Qy 67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
Db 5371 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 5430
Qy 127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
Db 5431 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 5490
Qy 187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
Db 5491 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 5550
Qy 247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
Db 5551 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 5610
Qy 307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
Db 5611 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 5670
Qy 367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
Db 5671 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 5730
Qy 427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
Db 5731 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 5790
Qy 487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
Db 5791 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 5850
Qy 547 TCTGGCCCACCGACCAA---------GCCCTCCATGCCCTACCTGCTGAACAACAGGACT 597
Db 5851 TCTGGCCCACCGACCAAGCCCTCCATGCCCTCCATGCCCTACCTGCTGAACAACAGGACT 5910
Qy 598 TCCTGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATAC 657
Db 5911 TCCTGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATAC 5970
Qy 658 GAGATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAG 717
Db 5971 GAGATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAG 6030
Qy 718 GTTCAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATG 777
Db 6031 GTTCAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATG 6090
Qy 778 GCCAAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCA 837
Db 6091 GCCAAACCTACAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCA 6150
Qy 838 TTGACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCG 897
Db 6151 TTGACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCG 6210
Qy 898 ATGCCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAAC 957
Db 6211 ATGCCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAAC 6270
Qy 958 CAAAGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCT 1017
Db 6271 CAAAGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCT 6330
Qy 1018 GCCGGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGC 1077
Db 6331 GCCGGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGC 6390
Qy 1078 GAGACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCC 1137
Db 6391 GAGACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCC 6450
Qy 1138 TTGATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGT 1197
Db 6451 TTGATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGT 6510
Qy 1198 GTGAGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACC 1257
Db 6511 GTGAGATGTGCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACC 6570
Qy 1258 ACGGACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGA 1317
Db 6571 ACGGACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGA 6630
Qy 1318 CAGGCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTG 1377
Db 6631 CAGGCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTG 6690
Qy 1378 CTCATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACG 1437
Db 6691 CTCATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACG 6750
Qy 1438 GAATCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGG 1497
Db 6751 GAATCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGG 6810
Qy 1498 CCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGG 1557
Db 6811 CCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGG 6870
Qy 1558 ACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACG 1617
Db 6871 ACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACG 6930
Qy 1618 AGCACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACT 1677
Db 6931 AGCACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACT 6990
Qy 1678 ATGTCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGG 1737
Db 6991 ATGTCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGG 7050
Qy 1738 ACAATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTG 1797
Db 7051 ACAATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTG 7110
Qy 1798 TTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCA 1857
Db 7111 TTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCA 7170
Qy 1858 GAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCC 1917
Db 7171 GAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCC 7230
Qy 1918 AGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACC 1977
Db 7231 AGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACC 7290
Qy 1978 CCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGAC 2037
Db 7291 CCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGAC 7350
Qy 2038 CTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACT 2097
Db 7351 CTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAA---------- 7400
Qy 2098 GCACCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGG 2157
Db 7401 ---------- 7400
Qy 2158 TCCTGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACA 2217

```
Db       7401 ------------------------------------GGAAGTGCTGGCCTATTCCAACA 7423

Qy       2218 GCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7424 GCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCC 7483

Qy       2278 TCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCG 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7484 TCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCG 7543

Qy       2338 AGCACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCC 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7544 AGCACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCC 7603

Qy       2398 TGCAAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGA 2457
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7604 TGCAAACGAGGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGA 7663

Qy       2458 CGGAGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATG 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7664 CGGAGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATG 7723

Qy       2518 GGATCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7724 GGATCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCC 7783

Qy       2578 ACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCT 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7784 ACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCT 7843

Qy       2638 TGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGT 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7844 TGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGT 7903

Qy       2698 CGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7904 CGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACC 7963

Qy       2758 CCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACT 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7964 CCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACT 8023

Qy       2818 CTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGG 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8024 CTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGG 8083

Qy       2878 AGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTT 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8084 AGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTT 8143

Qy       2938 GCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 2997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8144 GCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 8203

Qy       2998 ATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCC 3057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8204 ATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCC 8263

Qy       3058 TCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTG 3117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8264 TCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTG 8323

Qy       3118 CCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTG 3177
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       8324 CCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTG 8383

Qy       3178 TATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAA 3237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8384 TATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAA 8443

Qy       3238 CAGAAAAAAAAAAAAAAAAAAA 3260
             |||||| ||| |||||||||||

Db       8444 CAGAAACAAAGTCAAAAAAAAAA 8466
```

```
RESULT 9
US-10-133-172-19
; Sequence 19, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
;  APPLICANT: WEIGEL, PAUL H
;  APPLICANT:  WEIGEL, JANET A
;  TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
;  FILE REFERENCE: 5864.014
;  CURRENT APPLICATION NUMBER: US/10/133,172
;  CURRENT FILING DATE:  2002-04-25
;  PRIOR APPLICATION NUMBER: 60/286,468
;  PRIOR FILING DATE: 2001-04-25
;  PRIOR APPLICATION NUMBER: 09/842,930
;  PRIOR FILING DATE: 2001-04-25
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
;   LENGTH: 4962
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-133-172-19

  Query Match             87.5%;  Score 2853.2;  DB 15;  Length 4962;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2855;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2105 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 2164

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2165 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 2224

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2225 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 2284

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2285 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 2344

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2345 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 2404

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2405 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 2464

Qy        367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2465 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 2524

Qy        427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2525 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 2584

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2585 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 2644

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2645 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 2704

Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2705 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 2764
```

```
Qy       667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
Db      2765 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 2824

Qy       727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
Db      2825 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 2884

Qy       787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
Db      2885 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 2944

Qy       847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
Db      2945 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 3004

Qy       907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
Db      3005 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 3064

Qy       967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
Db      3065 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 3124

Qy      1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
Db      3125 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 3184

Qy      1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
Db      3185 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 3244

Qy      1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
Db      3245 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 3304

Qy      1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
Db      3305 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 3364

Qy      1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
Db      3365 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 3424

Qy      1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
Db      3425 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 3484

Qy      1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
Db      3485 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 3544

Qy      1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
Db      3545 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 3604

Qy      1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
Db      3605 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 3664

Qy      1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
Db      3665 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 3724

Qy      1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
Db      3725 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 3784

Qy      1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
Db      3785 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 3844

Qy      1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
Db      3845 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 3904

Qy      1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
Db      3905 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 3964

Qy      1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
Db      3965 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 4024

Qy      1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
Db      4025 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 4084

Qy      1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
Db      4085 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 4144

Qy      2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
Db      4145 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 4204

Qy      2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
Db      4205 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 4264

Qy      2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
Db      4265 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 4324

Qy      2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
Db      4325 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 4384

Qy      2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
Db      4385 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 4444

Qy      2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
Db      4445 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 4504

Qy      2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
Db      4505 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 4564

Qy      2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
Db      4565 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 4624

Qy      2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
Db      4625 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 4684

Qy      2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
Db      4685 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 4744

Qy      2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
Db      4745 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 4804

Qy      2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
Db      4805 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 4864

Qy      2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
Db      4865 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 4924

Qy      2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGA 2864
```

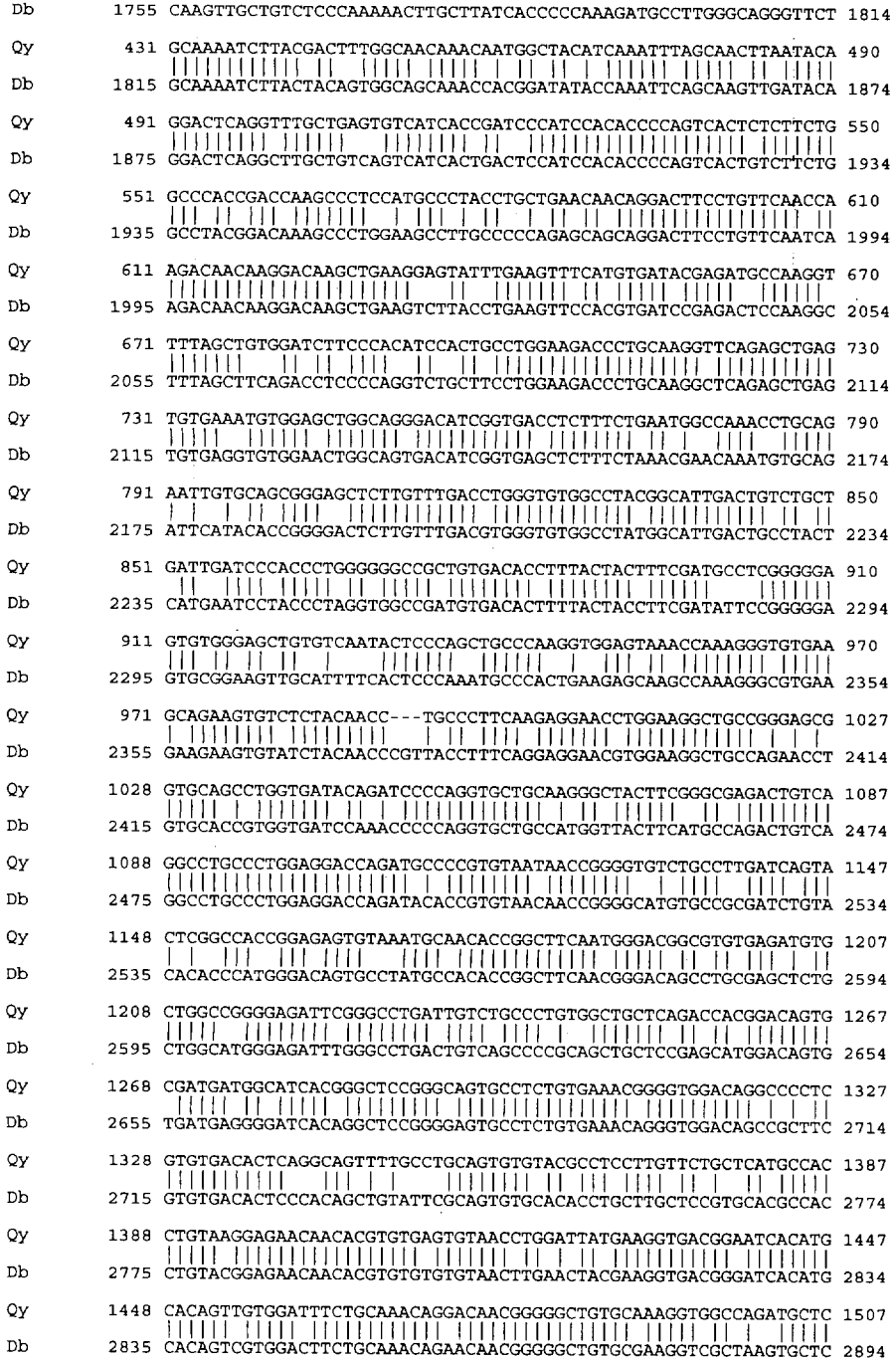
```
Db       4925 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGA 4962

RESULT 10
US-09-842-930A-1
; Sequence 1, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
;  APPLICANT: Weigel, Paul
;  TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
;  FILE REFERENCE: 5820.603
;  CURRENT APPLICATION NUMBER: US/09/842,930A
;  CURRENT FILING DATE:  2001-04-22
;  PRIOR APPLICATION NUMBER: 60/245,320
;  PRIOR FILING DATE: 2000-11-02
;  PRIOR APPLICATION NUMBER: 60/199,538
;  PRIOR FILING DATE: 2000-04-25
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 4706
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-09-842-930A-1

  Query Match             59.4%;  Score 1936.8;  DB 9;  Length 4706;
  Best Local Similarity   78.3%;  Pred. No. 0;
  Matches 2402;  Conservative    0;  Mismatches  652;  Indels   14;  Gaps    6;

Qy         11 GGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATTTCGT 70
Db       1395 GGAGCTTCCCAAGAACCCTTCGACGTCCCAGTACTTCTTCCAGTTGCAGGAGCATGCTGT 1454

Qy         71 GAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTGCACCTTTATCTGCAGCCTTTGA 130
Db       1455 CCGAGAGCTTGCTGGACCTGGCCCCTTCACCGTGTTCGCGCCTTTGTCTAGCTCCTTCAA 1514

Qy        131 TGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTCGGTA 190
Db       1515 TCATGAGCCCCGGATTAAAGACTGGGATCAGCAGGGCCTCATGTCCCAGGTTCTTCGCTA 1574

Qy        191 CCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATGCTAC 250
Db       1575 TCACGTGGTGGGCTGCCAGCAGCTGCTGTTGGACAACCTAAAAGTGACCACAAGTGCCAC 1634

Qy        251 TTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAAATAA 310
Db       1635 GACCCTCCAAGGAGAGCCAGTTTCCATCTCTGTCTCTCAGGACACTGTGTTCATAAACAA 1694

Qy        311 TAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCATAGA 370
Db       1695 TGAGGCGAAGGTCCTGTCCAGTGACATCATCAGCACCAATGGCGTCATCCACGTTATAGA 1754

Qy        371 CAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAATTCT 430
Db       1755 CAAGTTGCTGTCTCCCAAAAACTTGCTTATCACCCCCAAAGATGCCTTGGGCAGGGTTCT 1814

Qy        431 GCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAATACA 490
Db       1815 GCAAAATCTTACTACAGTGGCAGCAAACCACGGATATACCAAATTCAGCAAGTTGATACA 1874

Qy        491 GGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCTTCTG 550
Db       1875 GGACTCAGGCTTGCTGTCAGTCATCACTGACTCCATCCACACCCCAGTCACTGTCTTCTG 1934

Qy        551 GCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCAACCA 610
Db       1935 GCCTACGGACAAAGCCCTGGAAGCCTTGCCCCCAGAGCAGCAGGACTTCCTGTTCAATCA 1994

Qy        611 AGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCAAGGT 670
Db       1995 AGACAACAAGGACAAGCTGAAGTCTTACCTGAAGTTCCACGTGATCCGAGACTCCAAGGC 2054

Qy        671 TTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGCTGAG 730
Db       2055 TTTAGCTTCAGACCTCCCCAGGTCTGCTTCCTGGAAGACCCTGCAAGGCTCAGAGCTGAG 2114

Qy        731 TGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCTGCAG 790
Db       2115 TGTGAGGTGTGGAACTGGCAGTGACATCGGTGAGCTCTTTCTAAACGAACAAATGTGCAG 2174

Qy        791 AATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTCTGCT 850
Db       2175 ATTCATACACCGGGGACTCTTGTTTGACGTGGGTGTGGCCTATGGCATTGACTGCCTACT 2234

Qy        851 GATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGGGGGA 910
Db       2235 CATGAATCCTACCCTAGGTGGCCGATGTGACACTTTTACTACCTTCGATATTCCGGGGGA 2294

Qy        911 GTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTGTGAA 970
Db       2295 GTGCGGAAGTTGCATTTTCACTCCCAAATGCCCACTGAAGAGCAAGCCAAAGGGCGTGAA 2354

Qy        971 GCAGAAGTGTCTCTACAACC---TGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGCG 1027
Db       2355 GAAGAAGTGTATCTACAACCCGTTACCTTTCAGGAGGAACGTGGAAGGCTGCCAGAACCT 2414

Qy       1028 GTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTCA 1087
Db       2415 GTGCACCGTGGTGATCCAAACCCCCAGGTGCTGCCATGGTTACTTCATGCCAGACTGTCA 2474

Qy       1088 GGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGTA 1147
Db       2475 GGCCTGCCCTGGAGGACCAGATACACCGTGTAACAACCGGGGCATGTGCCGCGATCTGTA 2534

Qy       1148 CTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGTG 1207
Db       2535 CACACCCATGGGACAGTGCCTATGCCACACCGGCTTCAACGGGACAGCCTGCGAGCTCTG 2594

Qy       1208 CTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGTG 1267
Db       2595 CTGGCATGGGAGATTTGGGCCTGACTGTCAGCCCCGCAGCTGCTCCGAGCATGGACAGTG 2654

Qy       1268 CGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCTC 1327
Db       2655 TGATGAGGGGATCACAGGCTCCGGGGAGTGCCTCTGTGAAACAGGGTGGACAGCCCGCTTC 2714

Qy       1328 GTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCAC 1387
Db       2715 GTGTGACACTCCCACAGCTGTATTCGCAGTGTGCACACCTGCTTGCTCCGTGCACGCCAC 2774

Qy       1388 CTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACATG 1447
Db       2775 CTGTACGGAGAACAACACGTGTGTGTGTAACTTGAACTACGAAGGTGACGGGATCACATG 2834

Qy       1448 CACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCTC 1507
Db       2835 CACAGTCGTGGACTTCTGCAAACAGAACAACGGGGGCTGTGCGAAGGTCGCTAAGTGCTC 2894

Qy       1508 CCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACAG 1567
Db       2895 CCAGAAAGGCACCCAAGTCTCTTGCAGCTGCAAGAAAGGCTACAAGGGGGATGGCTACAG 2954

Qy       1568 CTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCAC 1627
Db       2955 CTGCATAGAGATAGACCCCTGTGCAGACGGTGTCAACGGGGGATGCCATGAGCACGCCAC 3014

Qy       1628 CTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAGA 1687
Db       3015 CTGCAGGATGACGGGCCCAGGCAAGCATAAGTGTGAATGTAAAAGTCACTATGTCGGGGA 3074

Qy       1688 TGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGCA 1747
Db       3075 CGGAGTGGACTGTGAGCCTGAGCAGCTGCCGCTCGACCGTTGCTTACAGGACAACGGACA 3134
```

```
Qy       1748 GTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGTT 1807
Db       3135 GTGCCACCCAGATGCCAGCTGTGCAGACCTCTACTTCCAGGACACGACCGTAGGAGTATT 3194

Qy       1808 CCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCTG 1867
Db       3195 CCATCTACGCTCCCCACTGGGCCAGTACAAACTGACATTTGACAAAGCCAAAGAAGCCTG 3254

Qy       1868 TGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAA 1927
Db       3255 TGCCAAAGAAGCTGCGACCATAGCCACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAA 3314

Qy       1928 GTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCTT 1987
Db       3315 GTATCACCTGTGCTCGGCCGGCTGGCTGGAGAGTGGGCGGGTTGCCTACCCGACTACGTA 3374

Qy       1988 CGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCAA 2047
Db       3375 TGCCTCTCAGAAGTGTGGTGCAAACGTTGTTGGGATCGTAGACTACGGATCCAGGGCCAA 3434

Qy       2048 CAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCAA 2107
Db       3435 CAAGAGTGAAATGTGGGATGTCTTCTGTTACCGGATGAAAGATGTGAACTGCACCTGCAA 3494

Qy       2108 GGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGTC 2167
Db       3495 GGCAGGCTATGTGGGAGATGGCTTCTCGTGCAGTGGGAACCTGCTGCAGGTCCTCATGTC 3554

Qy       2168 CTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTCG 2227
Db       3555 CTTCCCCTCGCTCACAAACTTCCTGACAGAGGTGCTGGCTTTTTCCAAGAGCTCAGCCCG 3614

Qy       2228 AGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCC 2287
Db       3615 AGGACAGGCATTTTTGAAACACCTGACTGACCTGTCCATCCGTGGCACCCTGTTTGTGCC 3674

Qy       2288 ACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACCT 2347
Db       3675 ACAGAACAGTGGGCTACCGGGAAATAAGAGCCTGTCTGGCCGGGACATTGAGCACCACCT 3734

Qy       2348 CGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGAG 2407
Db       3735 CACTAATGTCAACGTCTCCTTTTACAATGACCTTGTCAATGGTACCTTTCTGAGGACTAT 3794

Qy       2408 GGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCAG 2467
Db       3795 GCTGGGAAGCCAACTGCTCATTACCTTCAGCCAGGACCAGCTCCACC---AAGAGACCAG 3851

Qy       2468 GTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCA 2527
Db       3852 GTTTGTGGATGGAAGATCCATTCTGCAGTGGGACATCATCGCCGCCAATGGAATCCTCCA 3911

Qy       2528 TGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCTT 2587
Db       3912 TATTATTTCTGAACCTTTGAGAGCTCCTCCCACGGCAGCAACGGCTGCCCACTCTGGCCT 3971

Qy       2588 GGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTTA 2647
Db       3972 GGGGACAGGTATATTCTGTGCCGTCGTCCTGGTCACTGGTGCGATTGCTCTGGCAGCTTA 4031

Qy       2648 CTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGA 2707
Db       4032 CTCTTACTTCCGGCTAAAGCAGCGAACCACTGGTTTCCAGCGTTTTGA-TCAGAAGAGGA 4090

Qy       2708 CATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGA 2767
Db       4091 CATTGATGTCTTGGCTTTTGGCAAGCAGCAGCCCAAGAATATCGCAAACCCTCTGTATGA 4150

Qy       2768 GAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACG 2827
Db       4151 GACCTCAGCGCCGGCACCCCCAGAGTCCTCCTGTGACCCCTTCACAGACCCTGGAGAACA 4210

Qy       2828 GCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGC 2887

Db       4211 GGATCTGGAGGACAGCGACCCTCTGGGGGCACTG---CGGTCCTGACATGAGAAGCCAGC 4267

Qy       2888 CATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGC---ACCAGTTGCCTTTT 2944
Db       4268 AAGCAACCACAGTCACGGTTCCACGGTGATTCCCAGCCCCAGCTGTCTCATGGATCAGTT 4327

Qy       2945 AGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGG 3004
Db       4328 GTTTTAAAGAATGACAACACTCATAAGCCAGCCATACCTCACCCTTCTGGTTAATCTGGG 4387

Qy       3005 GGTTGTTTCTGTGGGTGAGAGATGTGTTG-CTGTGCCCACCCAGTACAGCTTCCTCCTCT 3063
Db       4388 ATTGTCGCCAGGGCTAAGGAGCCATGTTGCCTGGATACCTGGGGGACCTCCACCTCCTCT 4447

Qy       3064 GACCCTTT 3071
Db       4448 GAGCCTAT 4455
```

RESULT 11
US-10-133-172-1
; Sequence 1, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4706
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-133-172-1

Query Match 59.4%; Score 1936.8; DB 15; Length 4706;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 652; Indels 14; Gaps 6;

```
Qy         11 GGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATTTCGT 70
Db       1395 GGAGCTTCCCAAGAACCCTTCGACGTCCCAGTACTTCTTCCAGTTGCAGGAGCATGCTGT 1454

Qy         71 GAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCTTTGA 130
Db       1455 CCGAGAGCTTGCTGGACCTGGCCCCTTCACCGTGTTCGCGCCTTTGTCTAGCTCCTTCAA 1514

Qy        131 TGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTCGGTA 190
Db       1515 TCATGAGCCCCGGATTAAAGACTGGGATCAGCAGGGCCTCATGTCCCAGGTTCTTCGCTA 1574

Qy        191 CCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATGCTAC 250
Db       1575 TCACGTGGTGGGCTGCCAGCAGCTGCTGTTGGACAACCTAAAAGTGACCACAAGTGCCAC 1634

Qy        251 TTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAAATAA 310
Db       1635 GACCCTCCAAGGAGAGCCAGTTTCCATCTCTGTCTCTCAGGACACTGTGTTCATAAACAA 1694

Qy        311 TAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCATAGA 370
Db       1695 TGAGGCGAAGGTCCTGTCCAGTGACATCATCAGCACCAATGGCGTCATCCACGTTATAGA 1754

Qy        371 CAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAATTCT 430
```


Db 1755 CAAGTTGCTGTCTCCCAAAAACTTGCTTATCACCCCCAAAGATGCCTTGGGCAGGGTTCT 1814
Qy 431 GCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAATACA 490
Db 1815 GCAAAATCTTACTACAGTGGCAGCAAACCACGGATATACCAAATTCAGCAAGTTGATACA 1874
Qy 491 GGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCTTCTG 550
Db 1875 GGACTCAGGCTTGCTGTCAGTCATCACTGACTCCATCCACACCCCAGTCACTGTCTTCTG 1934
Qy 551 GCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCAACCA 610
Db 1935 GCCTACGGACAAAGCCCTGGAAGCCTTGCCCCCAGAGCAGCAGGACTTCCTGTTCAATCA 1994
Qy 611 AGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCAAGGT 670
Db 1995 AGACAACAAGGACAAGCTGAAGTCTTACCTGAAGTTCCACGTGATCCGAGACTCCAAGGC 2054
Qy 671 TTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGCTGAG 730
Db 2055 TTTAGCTTCAGACCTCCCCAGGTCTGCTTCCTGGAAGACCCTGCAAGGCTCAGAGCTGAG 2114
Qy 731 TGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCTGCAG 790
Db 2115 TGTGAGGTGTGGAACTGGCAGTGACATCGGTGAGCTCTTTCTAAACGAACAAATGTGCAG 2174
Qy 791 AATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTCTGCT 850
Db 2175 ATTCATACACCGGGGACTCTTGTTTGACGTGGGTGTGGCCTATGGCATTGACTGCCTACT 2234
Qy 851 GATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGGGGGA 910
Db 2235 CATGAATCCTACCCTAGGTGGCCGATGTGACACTTTTACTACCTTCGATATTCCGGGGGA 2294
Qy 911 GTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTGTGAA 970
Db 2295 GTGCGGAAGTTGCATTTTCACTCCCAAATGCCCACTGAAGAGCAAGCCAAAGGGCGTGAA 2354
Qy 971 GCAGAAGTGTCTCTACAACC---TGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGCG 1027
Db 2355 GAAGAAGTGTATCTACAACCCGTTACCTTTCAGGAGGAACGTGGAAGGCTGCCAGAACCT 2414
Qy 1028 GTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTCA 1087
Db 2415 GTGCACCGTGGTGATCCAAACCCCCAGGTGCTGCCATGGTTACTTCATGCCAGACTGTCA 2474
Qy 1088 GGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGTA 1147
Db 2475 GGCCTGCCCTGGAGGACCAGATACACCGTGTAACAACCGGGGCATGTGCCGCGATCTGTA 2534
Qy 1148 CTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGTG 1207
Db 2535 CACACCCATGGGACAGTGCCTATGCCACACCGGCTTCAACGGGACAGCCTGCGAGCTCTG 2594
Qy 1208 CTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGTG 1267
Db 2595 CTGGCATGGGAGATTTGGGCCTGACTGTCAGCCCCGCAGCTGCTCCGAGCATGGACAGTG 2654
Qy 1268 CGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCTC 1327
Db 2655 TGATGAGGGGATCACAGGCTCCGGGGAGTGCCTCTGTGAAACAGGGTGGACAGCCGCTTC 2714
Qy 1328 GTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCAC 1387
Db 2715 GTGTGACACTCCCACAGCTGTATTCGCAGTGTGCACACCTGCTTGCTCCGTGCACGCCAC 2774
Qy 1388 CTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACATG 1447
Db 2775 CTGTACGGAGAACAACACGTGTGTGTGTAACTTGAACTACGAAGGTGACGGGATCACATG 2834
Qy 1448 CACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCTC 1507
Db 2835 CACAGTCGTGGACTTCTGCAAACAGAACAACGGGGGCTGTGCGAAGGTCGCTAAGTGCTC 2894
Qy 1508 CCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACAG 1567
Db 2895 CCAGAAAGGCACCCAAGTCTCTTGCAGCTGCAAGAAAGGCTACAAGGGGGATGGCTACAG 2954
Qy 1568 CTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCAC 1627
Db 2955 CTGCATAGAGATAGACCCCTGTGCAGACGGTGTCAACGGGGGATGCCATGAGCACGCCAC 3014
Qy 1628 CTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAGA 1687
Db 3015 CTGCAGGATGACGGGCCCAGGCAAGCATAAGTGTGAATGTAAAAGTCACTATGTCGGGGA 3074
Qy 1688 TGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGCA 1747
Db 3075 CGGAGTGGACTGTGAGCCTGAGCAGCTGCCGCTCGACCGTTGCTTACAGGACAACGGACA 3134
Qy 1748 GTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGTT 1807
Db 3135 GTGCCACCCAGATGCCAGCTGTGCAGACCTCTACTTCCAGGACACGACCGTAGGAGTATT 3194
Qy 1808 CCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCTG 1867
Db 3195 CCATCTACGCTCCCCACTGGGCCAGTACAAACTGACATTTGACAAAGCCAAAGAAGCCTG 3254
Qy 1868 TGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAA 1927
Db 3255 TGCCAAAGAAGCTGCGACCATAGCCACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAA 3314
Qy 1928 GTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCTT 1987
Db 3315 GTATCACCTGTGCTCGGCCGGCTGGCTGGAGAGTGGGCGGGTTGCCTACCCGACTACGTA 3374
Qy 1988 CGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCAA 2047
Db 3375 TGCCTCTCAGAAGTGTGGTGCAAACGTTGTTGGGATCGTAGACTACGGATCCAGGGCCAA 3434
Qy 2048 CAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCAA 2107
Db 3435 CAAGAGTGAAATGTGGGATGTCTTCTGTTACCGGATGAAAGATGTGAACTGCACCTGCAA 3494
Qy 2108 GGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGTC 2167
Db 3495 GGCAGGCTATGTGGGAGATGGCTTCTCGTGCAGTGGGAACCTGCTGCAGGTCCTCATGTC 3554
Qy 2168 CTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTCG 2227
Db 3555 CTTCCCCTCGCTCACAAACTTCCTGACAGAGGTGCTGGCTTTTTCCAAGAGCTCAGCCCG 3614
Qy 2228 AGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCC 2287
Db 3615 AGGACAGGCATTTTTGAAACACCTGACTGACCTGTCCATCCGTGGCACCCTGTTTGTGCC 3674
Qy 2288 ACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACCT 2347
Db 3675 ACAGAACAGTGGGCTACCGGGAAATAAGAGCCTGTCTGGCCGGGACATTGAGCACCACCT 3734
Qy 2348 CGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGAG 2407
Db 3735 CACTAATGTCAACGTCTCCTTTTACAATGACCTTGTCAATGGTACCTTTCTGAGGACTAT 3794
Qy 2408 GGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCAG 2467
Db 3795 GCTGGGAAGCCAACTGCTCATTACCTTCAGCCAGGACCAGCTCCACC---AAGAGACCAG 3851
Qy 2468 GTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCA 2527
Db 3852 GTTTGTGGATGGAAGATCCATTCTGCAGTGGGACATCATCGCCGCCAATGGAATCCTCCA 3911
Qy 2528 TGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCTT 2587
Db 3912 TATTATTTCTGAACCTTTGAGAGCTCCTCCCACGGCAGCAACGGCTGCCCACTCTGGCCT 3971

```
Qy       2588 GGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTTA 2647
Db       3972 GGGGACAGGTATATTCTGTGCCGTCGTCCTGGTCACTGGTGCGATTGCTCTGGCAGCTTA 4031

Qy       2648 CTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGA 2707
Db       4032 CTCTTACTTCCGGCTAAAGCAGCGAACCACTGGTTTCCAGCGTTTTGA-TCAGAAGAGGA 4090

Qy       2708 CATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGA 2767
Db       4091 CATTGATGTCTTGGCTTTTGGCAAGCAGCAGCCCAAGAATATCGCAAACCCTCTGTATGA 4150

Qy       2768 GAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACG 2827
Db       4151 GACCTCAGCGCCGGCACCCCCAGAGTCCTCCTGTGACCCCTTCACAGACCCTGGAGAACA 4210

Qy       2828 GCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGC 2887
Db       4211 GGATCTGGAGGACAGCGACCCTCTGGGGGCACTG---CGGTCCTGACATGAGAAGCCAGC 4267

Qy       2888 CATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGC---ACCAGTTGCCTTTT 2944
Db       4268 AAGCAACCACAGTCACGGTTCCACGGTGATTCCCAGCCCCAGCTGTCTCATGGATCAGTT 4327

Qy       2945 AGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGG 3004
Db       4328 GTTTTAAAGAATGACAACACTCATAAGCCAGCCATACCTCACCCTTCTGGTTAATCTGGG 4387

Qy       3005 GGTTGTTTCTGTGGGTGAGAGATGTGTTG-CTGTGCCCACCCAGTACAGCTTCCTCCTCT 3063
Db       4388 ATTGTCGCCAGGGCTAAGGAGCCATGTTGCCTGGATACCTGGGGGACCTCCACCTCCTCT 4447

Qy       3064 GACCCTTT 3071
Db       4448 GAGCCTAT 4455
```

RESULT 12
US-10-023-896-38
; Sequence 38, Application US/10023896
; Publication No. US20030027776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004P1
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 718
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-896-38

Query Match 21.7%; Score 708; DB 15; Length 718;
Best Local Similarity 99.9%; Pred. No. 9.3e-213;
Matches 708; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2552 ACCCCCTGCCCCCGTGACCTTGACCCACACTGGCTTGGGAGCAGGGATCTTCTTTGCCAT 2611
Db          9 ACCCCCTGCCCCCGTGACCTTGACCCACACTGGCTTGGGAGCAGGGATCTTCTTTGCCAT 68

Qy       2612 CATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTTACTCCTACTTTCGGATAAACCGGAG 2671
Db         69 CATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTTACTCCTACTTTCGGATAAACCGGAG 128

Qy       2672 AACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGACATTAATGTTGCAGCTCTTGGCAA 2731
Db         129 AACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGACATTAATGTTGCAGCTCTTGGCAA 188

Qy       2732 GCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGAGAGCACAACCTCAGCTCCCCCAGA 2791
Db         189 GCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGAGAGCACAACCTCAGCTCCCCCAGA 248

Qy       2792 ACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGCTTGAGGGCAATGACCCCTT 2851
Db         249 ACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGCTTGAGGGCAATGACCCCTT 308

Qy       2852 GAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGGGCCA 2911
Db         309 GAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGGGCCA 368

Qy       2912 TCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCA 2971
Db         369 TCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCA 428

Qy       2972 GAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGT 3031
Db         429 GAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGT 488

Qy       3032 TGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACT 3091
Db         489 TGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACT 548

Qy       3092 CTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTAACTGTGATCTTTCTTCCC 3151
Db         549 CTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTAACTGTGATCTTTCTTCCC 608

Qy       3152 TGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGCCCCTAGCCCAGTGCCTGACACAGGA 3211
Db         609 TGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGCCCCTAGCCCAGTGCCTGACACAGGA 668

Qy       3212 ACTGTGCACAATAAAGGTTTATGGAACAGAAAAAAAAAAAAAAAAAAAA 3260
Db         669 ACTGTGCACAATAAAGGTTTATGGAACAGAAAAAAAAAAAAAAAAAAAA 717
```

RESULT 13
US-09-822-830A-559
; Sequence 559, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 559
; LENGTH: 722
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-830A-559

Query Match 16.5%; Score 537; DB 9; Length 722;
Best Local Similarity 99.6%; Pred. No. 1.4e-158;

```
  Matches  548;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       2695 AGTCGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGA 2754
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 AGTCGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGA 231

Qy       2755 ACCCCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGG 2814
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 ACCCCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGG 291

Qy       2815 ACTCTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGAC 2874
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ACTCTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGAC 351

Qy       2875 GGGAGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCA 2934
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GGGAGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCA 411

Qy       2935 GTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGG 2994
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 GTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGG 471

Qy       2995 CTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCT 3054
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 CTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCT 531

Qy       3055 TCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATT 3114
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 TCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATT 591

Qy       3115 CTGCCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCT 3174
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        592 CTGCCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCT 651

Qy       3175 TTGTATCCCAG-CCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTAT 3233
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 TTGTATCCCAGCCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTAT 711

Qy       3234 GGAACAGAAA 3243
             ||||||||||
Db        712 GGAACAGAAA 721


RESULT 14
US-10-264-237-1320
; Sequence 1320, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA131P1
;  CURRENT APPLICATION NUMBER: US/10/264,237
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/16450
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: US 60/205,515
;  PRIOR FILING DATE: 2000-05-19
;  NUMBER OF SEQ ID NOS: 2876
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1320
;  LENGTH: 3681
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (5)..(5)
;  OTHER INFORMATION: n equals a,t,g, or c
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (7)..(7)
;  OTHER INFORMATION: n equals a,t,g, or c
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (9)..(9)
;  OTHER INFORMATION: n equals a,t,g, or c
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (31)..(31)
;  OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-1320

  Query Match             13.0%;  Score 424.6;  DB 16;  Length 3681;
  Best Local Similarity   50.6%;  Pred. No. 1.7e-122;
  Matches 1258;  Conservative    1;  Mismatches 1135;  Indels   93;  Gaps    6;

Qy        177 CAGGTTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTG 236
              | |||  |||| ||||| ||||| | ||| | | |||||      ||  |||||      |
Db        699 CTGGTGTTTCGCTACCACGTGGTTGGCTGTCGGCGGCTGCGGAGCGAGGACCTGCTGGAG 758

Qy        237 ATCTCAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACG 296
              |  || ||  |||||  |||  | ||| |   | ||  ||       || |||
Db        759 CAGGGGTACGCCACGGCCCTCTCAGGGCACCCACTGCGCTTCAGCGAGAGGGAGGGCAGC 818

Qy        297 GTGTATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATT 356
              | |||  | |||  |  ||     |  |  |||  ||           || ||  ||
Db        819 ATATACCTCAATGACTTCGCGCGCGTGGTGAGCAGCGACCATGAGGCCGTGAACGGCATC 878

Qy        357 GTTCATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAAC 416
              | ||  |||| |||    | ||    |||    |||   |          ||  |
Db        879 CTGCACTTCATTGACCGTGTCCTGCTGCCCCCCGAGGCGCTGCACTGGGAGCCTGATGAT 938

Qy        417 TCTGGAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTT 476
              ||   |       |  |||| |  ||  |    ||       |  || ||||  |  ||
Db        939 GCTCCCATCCCGAGGAGAAATGTCACCGCCGCCGCCCAGGGCTTCGGTTACAAGATCTTC 998

Qy        477 AGCAACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCA 536
              |||  | |  |  |||  | ||  | |||    |  |    ||  |   |||  |  ||
Db        999 AGCGGCCTCCTGAAGGTGGCCGGCCTCCTGCCCCTGCTTCGAGAGGCATCCCATAGGCCC 1058

Qy        537 GTCACTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGAC 596
              ||||  |  | ||||||||  |||   |||  |  |  || || |||  | ||  |  |||| |
Db       1059 TTCACAATGCTGTGGCCCACAGACGCCGCCTTTCGAGCTCTGCCTCCGGATCGCCAGGCC 1118

Qy        597 TTCCTGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATA 656
              |  |||| |  |  ||| ||   ||||||||     |    |  ||  |  ||  ||||
Db       1119 TGGCTGTACCATGAGGACCACCGTGACAAGCTAGCAGCCATTCTGCGGGGCCACATGATT 1178

Qy        657 CGAGATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAA 716
              ||  ||| | |||  || ||    || || ||||         |     ||| ||||
Db       1179 CGCAATGTCGAGGCCTTGGCATCTGACCTGCCCAACCTGGGCCCACTTCGAACCATGCAT 1238

Qy        717 GGTTCAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAAT 776
              ||  |    |  | |   ||  |    | | |      |||||| ||| |  ||  |
Db       1239 GGGACCCCCATCTCTTTCTCCTGCAG---CCGAACGCGGCCCGGTGAGCTCATGGTGGGT 1295

Qy        777 GGCCAAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGC 836
              |  |         | |||||||||||| |  |    |||||    ||  |||||| |||
Db       1296 GAGGATGATGCTCGCATTGTGCAGCGGCACTTGCCCTTTGAGGGTGGCCTGGCCTATGGC 1355

Qy        837 ATTGACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTA----- 891
              || |||   ||||||    ||  ||| || |  |||||||||  ||||   |
Db       1356 ATCGACCAGCTGCTGGAGCCACCTGGCCTTGGTGCTCGCTGTGACCACTTTGAGACCCGG 1415

Qy        892 ----------------------------------------------------CTTTCGATGCC 902
                                                                   ||   |  |
Db       1416 CCCCTGCGACTGAACACCTGCAGCATCTGTGGGCTGGAGCCACCCTGTCCTGAGGGGTCA 1475

Qy        903 TCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAG 962
              || | || |  |       | |    | |  || ||| | ||   | |   |
Db       1476 CAGGAGCAGGGCAGCCCTGAGGCCTGCTGGCGCTTCTACCCGAAGTTCTGGACGTCCCCT 1535

Qy        963 GGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGG 1022
              || |     | |  |  || | |   || |    | |||     ||  | |
Db       1536 CCGCTGCACTCTTTGGGATTACGCAGCGTCTGGGTCCACCCCAGCCTTTGGGGTAGGCCC 1595

Qy       1023 GAGCGGTGCAGCCTGGTGATACA----------------------GATCCCCAGGTGC 1058
```

```
Db       1596 CAAGGCCTGGGCAGGGGCTGCCACCGCAATTGTGTCACCACCACCTGGAAGCCCAGCTGC 1655

Qy       1059 TGCAAGGGCTACTTCGGGCGAGACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGT 1118

Db       1656 TGCCCTGGTCACTATGGCAGTGAGTGCCAAGCTTGCCCTGGCGGCCCCAGCAGCCCTTGT 1715

Qy       1119 AATAACCGGGGGTGTCTGCCTTGATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACC 1178

Db       1716 AGTGACCGTGGCGTGTGCATGGACGGCATGAGTGGCAGTGGGCAGTGTCTGTGCCGTTCA 1775

Qy       1179 GGCTTCAATGGGACGGCGTGTGAGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTG 1238

Db       1776 GGTTTTGCTGGGACAGCCTGTGAACTCTGTGCTCCTGGTGCCTTTGGGCCCCATTGTCAA 1835

Qy       1239 CCCTGTGGCTGCTCAGACCACGGACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGC 1298

Db       1836 GCCTGCCGCTGCACTGTGCATGGCCGCTGTGATGAGGGCCTTGGGGGCTCTGGCTCCTGC 1895

Qy       1299 CTCTGTGAAACGGGGTGGACAGGCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTG 1358

Db       1896 TTCTGTGATGAAGGCTGGACTGGGCCACGCTGTGAGGTGCAACTGGAGCTGCAGCCTGTG 1955

Qy       1359 TGTACGCCTCCTTGTTCTGCTCATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAAC 1418

Db       1956 TGTACCCCACCCTGTGCACCCGAGGCTGTGTGCCGTGCAGGCAACAGCTGTGAGTGCAGC 2015

Qy       1419 CTGGATTATGAAGGTGACGGAATCACATGCACAGTTGTGGATTTCTGCAAACAGGACAAC 1478

Db       2016 CTGGGCTATGAAGGGGATGGCCGYGTGTGTACAGTGGCAGACCTGTGCCAGGACGGGCAT 2075

Qy       1479 GGGGGCTGTGCAAAGGTGGCCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGC 1538

Db       2076 GGTGGCTGCAGTGAGCACGCCAACTGTAGCCAGGTAGGAACAATGGTCACTTGTACCTGC 2135

Qy       1539 CAGAAGGGATACAAAGGGGACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGC 1598

Db       2136 CTGCCCGACTACGAGGGTGATGGCTGGAGCTGCCGGGCCCGCAACCCCTGCACAGATGGC 2195

Qy       1599 CTTAACGGAGGGTGTCACGAGCACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAG 1658

Db       2196 CACCGCGGGGGCTGCAGCGAGCACGCCAACTGCTTGAGCACCGGCCTGAACACACGGCGC 2255

Qy       1659 TGTGAGTGTAAAAGTCACTATGTCGGAGATGGGCTGAACTGT---GAGCCGGAGCAGCTG 1715

Db       2256 TGTGAGTGCCACGCAGGCTACGTAGGCGATGGACTGCAGTGTCTGGAGGAGTCGGAACCA 2315

Qy       1716 CCCATTGACCGCTGCTTACAGGACAATGGGCAGTGCCATGCAGACGCCAAATGTGTCGAC 1775

Db       2316 CCTGTGGACCGCTGCTTGGGCCAGCCACCGCCCTGCCACTCAGATGCCATGTGCACTGAC 2375

Qy       1776 CTCCACTTCCAGGATACCACTGTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTAT 1835

Db       2376 CTGCACTTCCAGGAGAAACGGGCTGGCGTTTTCCACCTCCAGGCCACCAGCGGCCCTTAT 2435

Qy       1836 AAGCTGACCTTTGACAAAGCCAGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACC 1895

Db       2436 GGTCTGAACTTTTCGGAGGCTGAGGCGGCATGCGAAGCACAGGGAGCCGTCCTTGCTTCA 2495

Qy       1896 TACAACCAGCTCTCCTATGCCCAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTG 1955

Db       2496 TTCCCTCAGCTCTCTGCTGCCCAGCAGCTGGGCTTCCACCTGTGCCTCATGGGCTGGCTG 2555

Qy       1956 GAGACCGGGCGGGTTGCCTACCCCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTG 2015

Db       2556 GCCAATGGCTCCACTGCCCACCCTGTGGTTTTCCCTGTGGCGGACTGTGGCAATGGTCGG 2615

Qy       2016 GTTGGGATAGTGGACTATGGACCTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGC 2075

Db       2616 GTGGGCATAGTCAGCCTGGGTGCCCGCAAGAACCTCTCAGAACGCTGGGATGCCTACTGC 2675

Qy       2076 TATCGGATGAAAGATGTGAACTGCACCTGCAAGGTGGGCTATGTGGGAGATGG---CTTC 2132

Db       2676 TTCCGTGTGCAAGATGTGGCCTGCCGATGCCGAAATGGCTTCGTGGGTGACGGGATCAGC 2735

Qy       2133 TCATGCAGTGGGAACCTGCTGCAGGTCCTGATGTCCTTCCCCTCACTCACAAACTTCCTG 2192

Db       2736 ACGTGCAATGGGAAGCTGCTGGATGTGCTGGCTGCCACTGCCAACTTCTCCACCTTCTAT 2795

Qy       2193 ACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTG 2252

Db       2796 GGGATGCTATTGGGCTATGCCAATGCCACCCAGCGGGGTCTCGACTTCCTGGACTTCCTG 2855

Qy       2253 ACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAAT 2312

Db       2856 GATGATGAGCTCACGTATAAGACACTCTTCGTCCCTGTCAATGAAGGCTTTGTGGACAAC 2915

Qy       2313 GAGACCTTGTCTGGGCGGGACATCGAGCACCACCTCGCCAATGTCAGCATGTTTTTCTAC 2372

Db       2916 ATGACGCTGAGTGGCCCAGACTTGGAGCTGCATGCCTCCAACGCCACCCTCCTAAGTGCC 2975

Qy       2373 AATGACCTTGTCAATGGCACCACCCTGCAAACGAGGGTGGGAAGCAAGCTGCTCATCACT 2432

Db       2976 AACGCCAGCCAGGGGAAGTTGCTTCCGGCCCACTCAGGCCTCAGCCTCATCATCAGTGAC 3035

Qy       2433 GCCAGCCAGGAC------CCACTCCAACCGACGGAGACCAGGTTTGTTGATGGAAGAGCC 2486

Db       3036 GCAGGCCCTGACAACAGTTCCTGGGCCCCTGTGGCCCCAGGGACAGTTGTGGTTAGCCGT 3095

Qy       2487 ATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCATGTCATTTCCAGGCCTTTA 2546

Db       3096 ATCATTGTGTGGGACATCATGGCCTTCAATGGCATCATCCATGCTCTGGCCAGCCCCCTC 3155

Qy       2547 AAAGCACCCCCTGCCCCCGTGACCTTG 2573

Db       3156 CTGGCACCCCCACAGCCCCAGGCAGTG 3182
```

RESULT 15
US-09-918-995-30654
; Sequence 30654, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 30654
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(448)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30654

Query Match 12.5%; Score 408.8; DB 10; Length 448;
Best Local Similarity 98.1%; Pred. No. 4.2e-118;
Matches 413; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy       1186 ATGGGACGGCGTGTGAGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTG 1245

Db         28 ATCGGACGGCGCGNGAGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTG 87

Qy       1246 GCTGCTCAGACCACGGACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTG 1305

Db         88 GTTGCTCAAACCACGGACTGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTG 147

Qy       1306 AAACGGGGTGGACAGGCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGC 1365
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 AAACGGGGTGGACAGGCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGC 207

Qy       1366 CTCCTTGTTCTGCTCATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATT 1425
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 CTCCTTGTTCTGCTCATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATT 267

Qy       1426 ATGAAGGTGACGGAATCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCT 1485
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGAAGGTGACGGAATCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCT 327

Qy       1486 GTGCAAAGGTGGCCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGG 1545
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GTGCAAAGGTGGCCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGG 387

Qy       1546 GATACAAAGGGGACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACG 1605
                ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||| ||
Db        388 GATACAAAGGGGACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGGCTTACCG 447

Qy       1606 G 1606
                |
Db        448 G 448
```

Search completed: June 21, 2004, 16:49:26
Job time : 1332 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 10:15:26 ; Search time 12294 Seconds
(without alignments)
11493.272 Million cell updates/sec

Title: US-10-028-248A-210
Perfect score: 3260
Sequence: 1 ggcacgagcaggagcttccc.....aaaaaaaaaaaaaaaaaaaa 3260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3259	100.0	3260	6	AX686401	AX686401 Sequence
2	3259	100.0	3260	9	AF160476	AF160476 Homo sapi
3	3241.4	99.4	8266	9	HSA295695	AJ295695 Homo sapi
4	3234.2	99.2	8251	9	AY311388	AY311388 Homo sapi
5	3232.6	99.2	5604	9	AK074051	AK074051 Homo sapi
6	3231	99.1	4575	9	AK024503	AK024503 Homo sapi
7	3229.8	99.1	8214	9	AB052958	AB052958 Homo sapi
8	3094.2	94.9	4492	9	AY227444	AY227444 Homo sapi
9	3017.8	92.6	3642	9	HSM801377	AL133021 Homo sapi
10	2999.8	92.0	8444	6	AX686192	AX686192 Sequence
11	2995.6	91.9	8495	6	AX686194	AX686194 Sequence
12	2828.8	86.8	3625	6	AX149461	AX149461 Sequence
13	1949.2	59.8	8157	10	AF364951	AF364951 Mus muscu
14	1936.8	59.4	4706	10	AY007370	AY007370 Rattus no
15	1778.6	54.6	2483	6	AX149459	AX149459 Sequence
16	1777	54.5	2512	6	AX149455	AX149455 Sequence
17	1721.2	52.8	2011	6	AX149447	AX149447 Sequence
18	1429.6	43.9	1804	6	AX149449	AX149449 Sequence
19	1186	36.4	1500	4	BTA550060	AJ550060 Bos tauru
20	990.2	30.4	1624	6	AX149457	AX149457 Sequence
21	764.2	23.4	1577	6	AX149463	AX149463 Sequence
c 22	437.6	13.4	146283	2	AC022277	AC022277 Homo sapi
c 23	437.6	13.4	180064	9	AC025265	AC025265 Homo sapi
c 24	437.6	13.4	180525	2	AC063946	AC063946 Homo sapi
c 25	437.6	13.4	184435	2	AC141618	AC141618 Homo sapi
26	423.4	13.0	7870	9	HSA275213	AJ275213 Homo sapi
27	423.4	13.0	7909	9	D87433	D87433 Homo sapien
28	423.4	13.0	7916	9	AB052956	AB052956 Homo sapi
29	399.4	12.3	3911	10	BC031166	BC031166 Mus muscu
30	399.4	12.3	7931	10	AF290914	AF290914 Mus muscu
31	391.2	12.0	2863	6	AX747085	AX747085 Sequence
32	391.2	12.0	2863	9	AK091600	AK091600 Homo sapi
33	285	8.7	2160	6	AX748299	AX748299 Sequence
34	285	8.7	2160	9	AK093774	AK093774 Homo sapi
35	260	8.0	6824	6	AX781121	AX781121 Sequence
36	187.4	5.7	1482	6	AR269667	AR269667 Sequence
37	153.2	4.7	204736	2	AC113310	AC113310 Mus muscu
38	153.2	4.7	247224	2	AC112790	AC112790 Mus muscu
39	153.2	4.7	296666	2	AC112983	AC112983 Mus muscu
40	151.8	4.7	146283	2	AC022277	AC022277 Homo sapi
41	149.8	4.6	157452	2	AC010190	AC010190 Homo sapi
42	134.4	4.1	230880	2	AC025501	AC025501 Mus muscu
c 43	132.2	4.1	3526	9	AK126526	AK126526 Homo sapi
c 44	128.8	4.0	202968	2	AC137422	AC137422 Rattus no
c 45	124	3.8	142304	2	AC099529	AC099529 Felis cat

ALIGNMENTS

RESULT 1
AX686401
LOCUS AX686401 3260 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 210 from Patent WO02059315.
ACCESSION AX686401
VERSION AX686401.1 GI:29372140
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shimkets,R.A., Patturajan,M., Vernet,C.A., Casman,S.J.,
Malyankar,U., Shenoy,S., Spytek,K.A., Gangolli,E., Miller,C.,
Boldog,F., Li,L., Taupier,R.J., Kekuda,R., Smithson,G.,

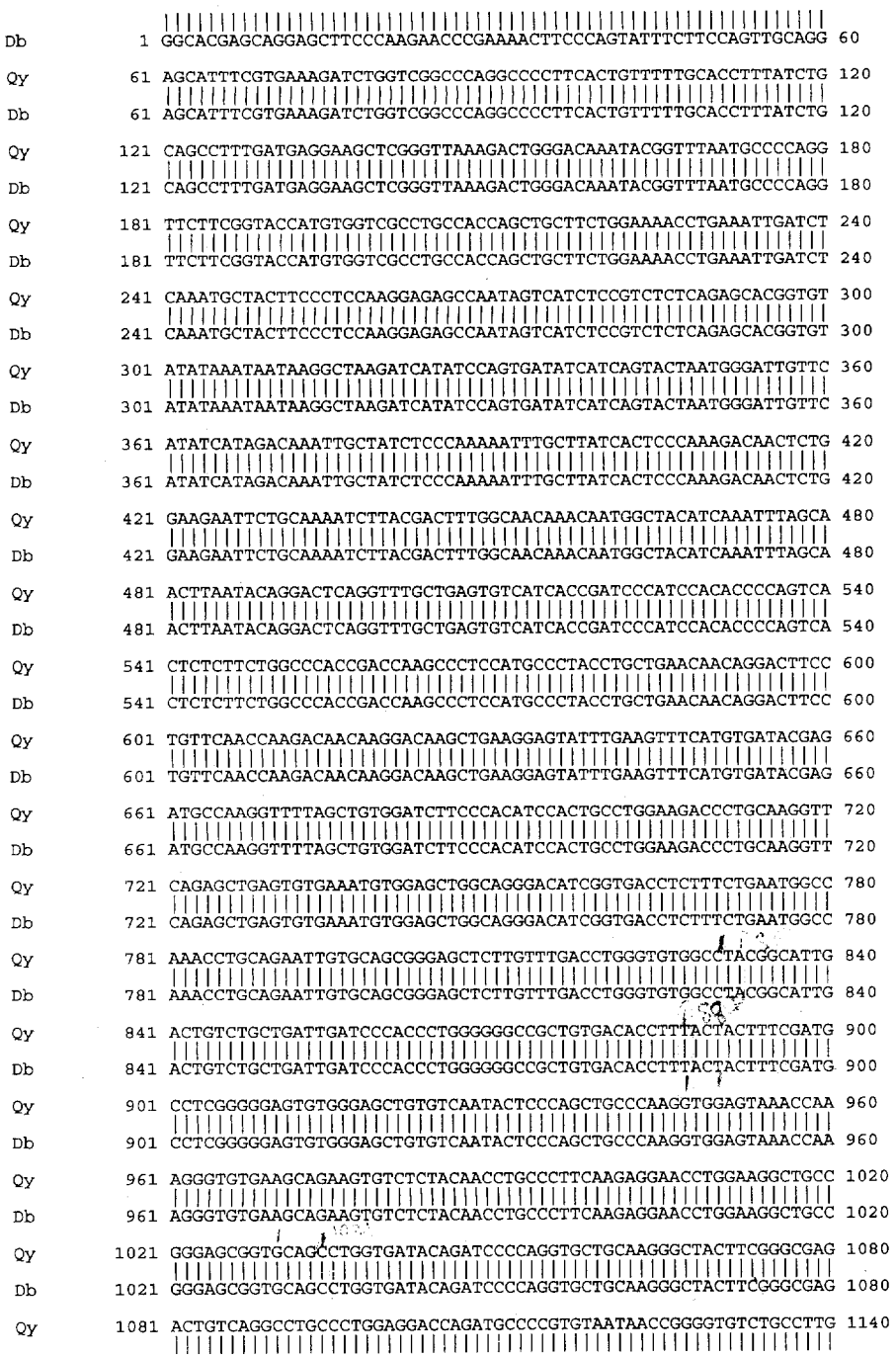
```
            Zerhusen,B.D., Liu,X., Colman,S.D., Tchernev,V., Si,J., Edinger,S.,
            Stone,D., Sciore,P., Millet,I. and Rothenberg,M.
  TITLE     Human nucleic acids and polypeptides and methods of use thereof
  JOURNAL   Patent: WO 02059315-A 210 01-AUG-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .3260
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           100.0%;  Score 3259;  DB 6;  Length 3260;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3260;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGG 60

Qy         61 AGCATTTCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCATTTCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTG 120

Qy        121 CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGG 180

Qy        181 TTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCT 240

Qy        241 CAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGT 300

Qy        301 ATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTC 360

Qy        361 ATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTG 420

Qy        421 GAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCA 480

Qy        481 ACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCA 540

Qy        541 CTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600

Qy        601 TGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAG 660

Qy        661 ATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTT 720

Qy        721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCC 780

Qy        781 AAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTG 840

Qy        841 ACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG 900

Qy        901 CCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAA 960

Qy        961 AGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCC 1020

Qy       1021 GGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAG 1080

Qy       1081 ACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTG 1140

Qy       1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG 1200

Qy       1201 AGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACG 1260

Qy       1261 GACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAG 1320

Qy       1321 GCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTC 1380

Qy       1381 ATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAA 1440

Qy       1441 TCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCA 1500

Qy       1501 GATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACG 1560

Qy       1561 GGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGGTGTCACGAGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGGTGTCACGAGC 1620

Qy       1621 ACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATG 1680

Qy       1681 TCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACA 1740

Qy       1741 ATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTG 1800

Qy       1801 GGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860

Qy       1861 AGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGA 1920

Qy       1921 AGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1921 AGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCA 1980
Qy       1981 CAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTA 2040
Db       1981 CAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTA 2040
Qy       2041 GACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCA 2100
Db       2041 GACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCA 2100
Qy       2101 CCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCC 2160
Db       2101 CCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCC 2160
Qy       2161 TGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCT 2220
Db       2161 TGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCT 2220
Qy       2221 CAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCT 2280
Db       2221 CAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCT 2280
Qy       2281 TTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGC 2340
Db       2281 TTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGC 2340
Qy       2341 ACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGC 2400
Db       2341 ACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGC 2400
Qy       2401 AAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGG 2460
Db       2401 AAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGG 2460
Qy       2461 AGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGA 2520
Db       2461 AGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGA 2520
Qy       2521 TCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACA 2580
Db       2521 TCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACA 2580
Qy       2581 CTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGG 2640
Db       2581 CTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGG 2640
Qy       2641 CTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGG 2700
Db       2641 CTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGG 2700
Qy       2701 AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCT 2760
Db       2701 AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCT 2760
Qy       2761 TGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTG 2820
Db       2761 TGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTG 2820
Qy       2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880
Db       2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880
Qy       2881 TGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCC 2940
Db       2881 TGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCC 2940
Qy       2941 TTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATC 3000
Db       2941 TTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATC 3000
Qy       3001 TGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCC 3060
Db       3001 TGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCC 3060
Qy       3061 TCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCC 3120
Db       3061 TCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCC 3120
Qy       3121 TACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTAT 3180
Db       3121 TACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTAT 3180
Qy       3181 CCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240
Db       3181 CCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240
Qy       3241 AAAAAAAAAAAAAAAAAAAA 3260
Db       3241 AAAAAAAAAAAAAAAAAAAA 3260
```

RESULT 2
AF160476

```
LOCUS       AF160476               3260 bp    mRNA    linear   PRI 12-JUL-2000
DEFINITION  Homo sapiens CD44-like precursor FELL mRNA, complete cds.
ACCESSION   AF160476
VERSION     AF160476.1  GI:9049505
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3260)
  AUTHORS   Tao,Q., Zhang,W. and Cao,X.
  TITLE     Molecular cloning and characterization of human FELL sharing
            homology with CD44
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3260)
  AUTHORS   Tao,Q., Zhang,W. and Cao,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-1999) Department of Immunology, Shanghai
            Brilliance Biotechnology Institute & Second Military Medical
            University, 800 Xiangyin Road, Shanghai 200433, P. R. China
FEATURES             Location/Qualifiers
     source          1. .3260
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             171. .2864
                     /codon_start=1
                     /product="CD44-like precursor FELL"
                     /protein_id="AAF82398.1"
                     /db_xref="GI:9049506"
                     /translation="MPQVLRYHVVACHQLLLENLKLISNATSLQGEPIVISVSQSTVY
                     INNKAKIISSDIISTNGIVHIIDKLLSPKNLLITPKDNSGRILQNLTTLATNNGYIKF
                     SNLIQDSGLLSVITDPIHTPVTLFWPTDQALHALPAEQQDFLFNQDNKDKLKEYLKFH
                     VIRDAKVLAVDLPTSTAWKTLQGSELSVKCGAGRDIGDLFLNGQTCRIVQRELLFDLG
                     VAYGIDCLLIDPTLGGRCDTFTTFDASGECGSCVNTPSCPRWSKPKGVKQKCLYNLPF
                     KRNLEGCRERCSLVIQIPRCCKGYFGRDCQACPGGPDAPCNNRGVCLDQYSATGECKC
                     NTGFNGTACEMCWPGRFGPDCLPCGCSDHGQCDDGITGSGQCLCETGWTGPSCDTQAV
                     LPAVCTPPCSAHATCKENNTCECNLDYEGDGITCTVVDFCKQDNGGCAKVARCSQKGT
                     KVSCSCQKGYKGDGHSCTEIDPCADGLNGGCHEHATCKMTGPGKHKCECKSHYVGDGL
                     NCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHLRSPLGQYKLTFDKAREAC
                     ANEAATMATYNQLSYAQKAKYHLCSAGWLETGRVAYPTAFASQNCGSGVVGIVDYGPR
                     PNKSEMWDVFCYRMKDVNCTCKVGYVGDGFSCSGNLLQVLMSFPSLTNFLTEVLAYSN
                     SSARGRAFLEHLTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLANVSMFFYNDLVNG
                     TTLQTRVGSKLLITASQDPLQPTETRFVDGRAILQWDIFASNGIIHVISRPLKAPPAP
                     VTLTHTGLGAGIFFAIILVTGAVALAAYSYFRINRRTIGFQHFESEEDINVAALGKQQ
                     PENISNPLYESTTSAPPEPSYDPFTDSEERQLEGNDPLRTL"
ORIGIN
```

```
  Query Match           100.0%;  Score 3259;  DB 9;  Length 3260;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3259;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGG 60
```


Db 1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGG 60
Qy 61 AGCATTTCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTG 120
Db 61 AGCATTTCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTG 120
Qy 121 CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGG 180
Db 121 CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGG 180
Qy 181 TTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCT 240
Db 181 TTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCT 240
Qy 241 CAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGT 300
Db 241 CAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGT 300
Qy 301 ATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTC 360
Db 301 ATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTC 360
Qy 361 ATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTG 420
Db 361 ATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTG 420
Qy 421 GAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCA 480
Db 421 GAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCA 480
Qy 481 ACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCA 540
Db 481 ACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCA 540
Qy 541 CTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600
Db 541 CTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600
Qy 601 TGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAG 660
Db 601 TGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAG 660
Qy 661 ATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTT 720
Db 661 ATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTT 720
Qy 721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCC 780
Db 721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCC 780
Qy 781 AAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTG 840
Db 781 AAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTG 840
Qy 841 ACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG 900
Db 841 ACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG 900
Qy 901 CCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAA 960
Db 901 CCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAA 960
Qy 961 AGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCC 1020
Db 961 AGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCC 1020
Qy 1021 GGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAG 1080
Db 1021 GGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAG 1080
Qy 1081 ACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTG 1140
Db 1081 ACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTG 1140
Qy 1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG 1200
Db 1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG 1200
Qy 1201 AGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACG 1260
Db 1201 AGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACG 1260
Qy 1261 GACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAG 1320
Db 1261 GACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAG 1320
Qy 1321 GCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTC 1380
Db 1321 GCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTC 1380
Qy 1381 ATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAA 1440
Db 1381 ATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAA 1440
Qy 1441 TCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCA 1500
Db 1441 TCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCA 1500
Qy 1501 GATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACG 1560
Db 1501 GATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACG 1560
Qy 1561 GGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGC 1620
Db 1561 GGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGC 1620
Qy 1621 ACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATG 1680
Db 1621 ACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATG 1680
Qy 1681 TCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACA 1740
Db 1681 TCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACA 1740
Qy 1741 ATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTG 1800
Db 1741 ATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTG 1800
Qy 1801 GGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860
Db 1801 GGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860
Qy 1861 AGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGA 1920
Db 1861 AGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGA 1920
Qy 1921 AGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCA 1980
Db 1921 AGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCA 1980
Qy 1981 CAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTA 2040
Db 1981 CAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTA 2040
Qy 2041 GACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCA 2100
Db 2041 GACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCA 2100
Qy 2101 CCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCC 2160
Db 2101 CCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCC 2160
Qy 2161 TGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCT 2220
Db 2161 TGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCT 2220

```
Qy       2221 CAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCT 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCT 2280

Qy       2281 TTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGC 2340

Qy       2341 ACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGC 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGC 2400

Qy       2401 AAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGG 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGG 2460

Qy       2461 AGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGA 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 AGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGA 2520

Qy       2521 TCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACA 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACA 2580

Qy       2581 CTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGG 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGG 2640

Qy       2641 CTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGG 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGG 2700

Qy       2701 AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCT 2760

Qy       2761 TGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTG 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTG 2820

Qy       2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880

Qy       2881 TGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCC 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCC 2940

Qy       2941 TTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATC 3000

Qy       3001 TGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCC 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCC 3060

Qy       3061 TCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCC 3120

Qy       3121 TACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTAT 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCWTCTTTGTAT 3180

Qy       3181 CCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240

Qy       3241 AAAAAAAAAAAAAAAAAAAA 3260
             ||||||||||||||||||||
Db       3241 AAAAAAAAAAAAAAAAAAAA 3260
```

```
RESULT 3
HSA295695
LOCUS       HSA295695              8266 bp    mRNA    linear   PRI 28-JAN-2003
DEFINITION  Homo sapiens mRNA for stabilin-2 (STAB2 gene).
ACCESSION   AJ295695
VERSION     AJ295695.1  GI:18073427
KEYWORDS    STAB2 gene; stabilin-2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Politz,O., Gratchev,A., McCourt,P.A., Schledzewski,K., Guillot,P.,
            Johansson,S., Svineng,G., Franke,P., Kannicht,C., Kzhyshkowska,J.,
            Longati,P., Velten,F.W., Johansson,S. and Goerdt,S.
  TITLE     Stabilin-1 and -2 constitute a novel family of fasciclin-like
            hyaluronan receptor homologues
  JOURNAL   Biochem. J. 362 (Pt 1), 155-164 (2002)
  MEDLINE   21818465
   PUBMED   11829752
REFERENCE   2
  AUTHORS   Politz,O., Gratchev,A., McCourt,P.A.G., Schledzewski,K.,
            Guillot,P., Johannson,S., Birk,R., Hakiy,N., Franke,P., Kodelja,V.,
            Kannicht,C., Orfanos,C.E., Johannson,S. and Goerdt,S.
  TITLE     Stabilin-1 and stabilin-2 constitute a novel family of fasciclin
            domain-containing adhesion molecules associated with
            endothelial-macrophage differentiation and angiogenic processes
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 8266)
  AUTHORS   Gratchev,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUL-2000) Gratchev A., Department of Dermatology,
            University Medical Center Mannheim, Theodor-Kutzer-Ufer 1, D-68135
            Mannheim, GERMANY
  REMARK    Revised by author 27-APR-2001
FEATURES             Location/Qualifiers
     source          1. .8266
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .8266
                     /gene="STAB2"
     CDS             187. .7842
                     /gene="STAB2"
                     /codon_start=1
                     /product="stabilin-2"
                     /protein_id="CAC82105.1"
                     /db_xref="GI:18073428"
                     /db_xref="GOA:Q8WWQ8"
                     /db_xref="SPTREMBL:Q8WWQ8"
                     /translation="MMLQHLVIFCLGLVVQNFCSPAETTGQARRCDRKSLLTIRTECR
                     SCALNLGVKCPDGYTMITSGSVGVRDCRYTFEVRTYSLSLPGCRHICRKDYLQPRCCP
                     GRWGPDCIECPGGAGSPCNGRGSCAEGMEGNGTCSCQEGFGGTACETCADDNLFGPSC
                     SSVCNCVHGVCNSGLDGDGTCECYSAYTGPKCDKPIPECAALLCPENSRCSPSTEDEN
                     KLECKCLPNYRGDGKYCDPINPCLRKICHPHAHCTYLGPNRHSCTCQEGYRGDGQVCL
                     PVDPCQINFGNCPTKSTVCKYDGPGQSHCECKEHYQNFVPGVGCSMTDICKSDNPCHR
                     NANCTTVAPGRTECICQKGYVGDGLTCYGNIMERLRELNTEPRGKWQGRLTSFISLLD
                     KAYAWPLSKLGPFTVLLPTDKGLKGFNVNELLVDNKAAQYFVKLHIIAGQMNIEYMNN
                     TDMFYTLTGKSGEIFNSDKDNQIKLKLHGGKKKVKIIQGDIIASNGLLHILDRAMDKL
                     EHTFESNNEQTIMTMLQPRYSKFRSLLEETNLGHALDEDGVGGPYTIFVPNNEALNNM
                     KDGTLDYLLSPEGSRKLLELVRYHIVPFTQLEVATLISTPHIRSMANQLIQFNTTDNG
                     QILANDVAMEEIEITAKNGRIYTLTGVLIPPSIVPILPHRCDETKREMKLGTCVSCSL
                     VYWSRCPANSEPTALFTHRCVYSGRFGSLKSGCARYCNATVKIPKCCKGFYGPDCNQC
                     PGGFSNPCSGNGQCADSLGGNGTCICEEGFQGSQCQFCSDPNKYGPRCNKKCLCVHGT
                     CNNRIDSDGACLTGTCRDGSAGRLCDKQTSACGPYVQFCHIHATCEYSNGTASCICKA
                     GYEGDGTLCSEMDPCTGLTPGGCSRNAECIKTGTGTHTCVCQQGWTGNGRDCSEINNC
                     LLPSAGGCHDNASCLYVGPGQNECECKKGFRGNGIDCEPITSCLEQTGKCHPLASCQS
                     TSSGVWSCVCQEGYEGDGFLCYGNAAVELSFLSEAAIFNRWINNASLQPTLSATSNLT
                     VLVPSQQATEDMDQDEKSFWLSQSNIPALIKYHMLLGTYRVADLQTLSSSDMLATSLQ
                     GNFLHLAKVDGNITIEGASIVDGDNAATNGVIHIINKVLVPQRRLTGSLPNLLMRLEQ
                     MPDYPIFRGYIIQYNLANAIEAADAYTVFAPNNNAIENYIREKKVLSLEEDVLRYHVV
                     LEEKLLKNDLHNGMHRETMLGFSYFLSFFLHNDQLYVNEAPINYTNVATDKGVIHGLG
```

KVLEIQKNRCDNNATTIIRGRCRTCSSELTCPFGTKSLGNEKRRCIYTSYFMGRRTLF
IGCQPKCVRTVITRECCAGFFGPQCQPCPGNAQNVCFGNGICLDGVNGTGVCECGEGF
SGTACETCTEGKYGIHCDQACSCVHGRCNQGPLGDGSCDCDVGWRGVHCDNATTEDNC
NGTCHTSANCLTNSDGTASCKCAAGFQGNGTICTAINACEISNGGCSAKADCKRTTPG
RRVCTCKAGYTGDGIVCLEINPCLENHGGCDKNAECTQTGPNQAACNCLPAYTGDGKV
CTLINVCLTKNGGCSEFAICNHTGQVERTCTCKPNYIGDGFTCRGSIYQELPKNPKTS
QYFFQLQEHFVKDLVGPGPFTVFAPLSAAFDEEARVKDWDKYGLMPQVLRYHVVACHQ
LLLENLKLISNATSLQGEPIVISVSQSTVYINNKAKIISSDIISTNGIVHIIDKLLSP
KNLLITPKDNSGRILQNLTTLATNNGYIKFSNLIQDSGLLSVITDPIHTPVTLFWPTD
QALHALPAEQQDFLFNQDNKDKLKEYLKFHVIRDAKVLAVDLPTSTAWKTLQGSELSV
KCGAGRDIGDLFLNGQTCRIVQRELLFDLGVAYGIDCLLIDPTLGGRCDTFTTFDASG
ECGSCVNTPSCPRWSKPKGVKQKCLYNLPFKRNLEGCRERCSLVIQIPRCCKGYFGRD
CQACPGGPDAPCNNRGVCLDQYSATGECKCNTGFNGTACEMCWPGRFGPDCLPCGCSD
HGQCDDGITGSGQCLCETGWTGPSCDTQAVLPAVCTPPCSAHATCKENNTCECNLDYE
GDGITCTVVDFCKQDNGGCAKVARCSQKGTKVSCSCQKGYKGDGHSCTEIDPCADGLN
GGCHEHATCKMTGPGKHKCECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDL
HFQDTTVGVFHLRSPLGQYKLTFDKAREACANEAATMATYNQLSYAQKAKYHLCSAGW
LETGRVAYPTAFASQNCGSGVVGIVDYGPRPNKSEMWDVFCYRMKDVNCTCKVGYVGD
GFSCSGNLLQVLMSFPSLTNFLTEVLAYSNSSARGRAFLEHLTDLSIRGTLFVPQNSG
LGENETLSGRDIEHHLANVSMFFYNDLVNGTTLQTRLGSKLLITASQDPLQPTETRFV
DGRAILQWDIFASNGIIHVISRPLKAPPAPVTLTHTGLGAGIFFAIILVTGAVALAAY
SYFRINRRTIGFQHFESEEDINVAALGKQQPENISNPLYESTTSAPPEPSYDPFTDSE
ERQLEGNDPLRTL"

ORIGIN

Query Match 99.4%; Score 3241.4; DB 9; Length 8266;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3245; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```
Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4985 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5044

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5045 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 5104

Qy         127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5105 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 5164

Qy         187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5165 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 5224

Qy         247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5225 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 5284

Qy         307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5285 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 5344

Qy         367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5345 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 5404

Qy         427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5405 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 5464

Qy         487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5465 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 5524

Qy         547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5525 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 5584

Qy         607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5585 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 5644

Qy         667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5645 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 5704

Qy         727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5705 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 5764

Qy         787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5765 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 5824

Qy         847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5825 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 5884

Qy         907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5885 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 5944

Qy         967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5945 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 6004

Qy         1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6005 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 6064

Qy         1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6065 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 6124

Qy         1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6125 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 6184

Qy         1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
                ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       6185 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 6244

Qy         1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6245 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 6304

Qy         1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6305 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 6364

Qy         1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6365 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 6424

Qy         1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6425 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 6484

Qy         1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6485 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 6544

Qy         1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6545 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 6604

Qy         1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6605 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 6664

Qy         1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6665 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 6724

Qy         1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
```

```
Db       6725 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 6784

Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866

Db       6785 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 6844

Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926

Db       6845 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 6904

Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986

Db       6905 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 6964

Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046

Db       6965 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 7024

Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106

Db       7025 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 7084

Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166

Db       7085 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 7144

Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226

Db       7145 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 7204

Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286

Db       7205 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 7264

Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346

Db       7265 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 7324

Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406

Db       7325 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 7384

Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466

Db       7385 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 7444

Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526

Db       7445 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 7504

Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586

Db       7505 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 7564

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646

Db       7565 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 7624

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706

Db       7625 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 7684

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766

Db       7685 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 7744

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826

Db       7745 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 7804

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886

Db       7805 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 7864

Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946

Db       7865 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 7924

Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006

Db       7925 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 7984

Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066

Db       7985 TTGTTTCTGTGGGTGAGARATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 8044

Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126

Db       8045 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 8104

Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186

Db       8105 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 8164

Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246

Db       8165 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 8224

Qy       3247 AAAAAAAAAAAAAA 3260

Db       8225 AGTCAAAAAAAAAA 8238
```

```
RESULT 4
AY311388
LOCUS       AY311388               8251 bp    mRNA    linear   PRI 30-JUN-2003
DEFINITION  Homo sapiens FEX2 mRNA, complete cds.
ACCESSION   AY311388
VERSION     AY311388.1  GI:32351284
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 8251)
  AUTHORS   Park,S.-Y. and Kim,I.-S.
  TITLE     FEX2, a novel cell adhesion molecule of Fas-1 superfamily mediates
            cell-cell interaction
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 8251)
  AUTHORS   Park,S.-Y. and Kim,I.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAY-2003) Department of Biochemistry, School of
            Medicine, Kyungpook National University, 101 Dongin-dong, Jung-gu,
            Deagu 700-422, Korea
FEATURES             Location/Qualifiers
     source          1. .8251
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12q23.3"
     CDS             205. .7860
                     /note="FAS1, EGF-like and X-link domain containing
                     adhesion molecule-2"
                     /codon_start=1
                     /product="FEX2"
                     /protein_id="AAP74958.1"
                     /db_xref="GI:32351285"
                     /translation="MMLQHLVIFCLGLVVQNFCSPAETTGQARRCDRKSLLTIRTECR
                     SCALNLGVKCPDGYTMITSGSVGVRDCRYTFEVRTYSLSLPGCRHICRKDYLQPRCCP
                     GRWGPDCIECPGGAGSPCNGRGSCAEGMEGNGTCSCQEGFGGTACETCADDNLFGPSC
                     SSVCNCVHGVCNSGLDGDGTCECYSAYTGPKCDKPIPECAALLCPENSRCSPSTEDEN
                     KLECKCLPNYRGDGKYCDPINPCLRKICHPHAHCTYLGPNRHSCTCQEGYRGDGQVCL
                     PVDPCQINFGNCPTKSTVCKYDGPGQSHCECKEHYQNFVPGVGCSMTDICKSDNPCHR
```

NANCTTVAPGRTECICQKGYVGDGLTCYGNIMERLRELNTEPRGKWQGRLTSFISLLD
KAYAWPLSKLGPFTVLLPTDKGLKGFNVNELLVDNKAAQYFVKLHIIAQMNIEYMNN
TDMFYTLTGKSGEIFNSDKDNQIKLKLHGGKKKVKIIQGDIIASNGLLHILDRAMDKL
EPTFESNNEQTIMTMLQPRYSKFRSLLEETNLGHALDEDGVGGPYTIFVPNNEALNNM
KDGTLDYLLSPEGSRKLLELVRYHIVPFTQLEVATLISTPHIRSMANQLIQFNTTDNG
QILANDVAMEEIEITAKNGRIYTLTGVLIPPSIVPILPHRCDETKREMKLGTCVSCSL
VYWSRCPANSEPTALFTHRCVYSGRFGSLKSGCARYCNATVKIPKCCKGFYGPDCNQC
PGGFSNPCSGNGQCADSLGGNGTCICEEGFQGSQCQFCSDPNKYGPRCNKKCLCVHGT
CNNRIDSDGACLTGTCRDGSAGRLCDKQTSACGPYVQFCHIHATCEYSNGTASCICKA
GYEGDGTLCSEMDPCTGLTPGGCSRNAECIKTGTGTHTCVCQQGWTGNGRDCSEINNC
LLPSAGGCHDNASCLYVGPGQNECECKKGFRGNGIDCEPITSCLEQTGKCHPLASCQS
TSSGVWSCVCQEGYEGDGFLCYGNAAVELSFLSEAAIFNRWINNASLQPTLSATSNLT
VLVPSQQATEDMDQDEKSFWLSQSNIPALIKYHMLLGTYRVADLQTLSSSDMLATSLQ
GNFLHLAKVDGNITIEGASIVDGDNAATNGVIHIIINKVLVPQRRLTGSLPNLLMRLEQ
MPDYSIFRGYIIQYNLANAIEAADAYTVFAPNNNAIENYIREKKVLSLEEDVLRYHVV
LEEKLLKNDLHNGMHRETMLGFSYFLSFFLHNDQLYVNEAPINYTNVATDKGVIHGLG
KVLEIQKNRCDNNDTTIIRGRCRTCSSELTCPFGTKSLGNEKRRCIYTSYFMGRRTLF
IGCQPKCVRTVITRECCAGFFGPQCQPCPGNAQNVCFGNGICLDGVNGTGVCECGEGF
SGTACETCTEGKYGIHCDQACSCVHGRCNQGPLGDGSCDCDVGWRGVHCDNATTEDNC
NGTCHTSANCLTNSDGTASCKCAAGFQGNGTICTAINACEISNGGCSAKADCKRTTPG
RRVCTCKAGYTGDGIVCLEINPCLENHGGCDKNAECTQTGPNQAACNCLPAYTGDGKV
CTLINVCLTKNGGCSEFAICNHTGQVERTCTCKPNYIGDGFTCRGSIYQELPKNPKTS
QYFFQLQEHFVKDLVGPGPFTVFAPLSAAFDEEARVKDWDKYGLMPQVLRYHVVACHQ
LLLENLKLISNATSLQGEPIVISVSQSTVYINNKAKIISSDIISTNGIVHIIDKLLSP
KNLLITPKDNSGRILQNLTTLATNNGYIKFSNLIQDSGLLSVITDPIHTPVTLFWPTD
QALHALPAEQQDFLFNQDNKDKLKEYLKFHVIRDAKVLAVDLPTSTAWKTLQGSELSV
KCGAGRDIGDLFLNGQTCRIVQRELLFDLGVAYGIDCLLIDPTLGGRCDTFTTFDASG
ECGSCVNTPSCPRWSKPKGVKQKCLYNLPFKRNLEGCRERCSLVIQIPRCCKGYFGRD
CQACPGGPDAPCNNRGVCLDQYSATGECKCNTGFNGTACEMCWPGRFGPDCLPCGCSD
HGQCDDGITGSGQCLCETGWTGPSCDTQAVLPAVCTPPCSAHATCKENNTCECNLDYE
GDGITCTVVDFCKQDNGGCAKVARCSQKGTKVSCSCQKGYKGDGHSCTEIDPCADGLN
GGCHEHATCKMTGPGKHKCECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDL
HFQDTTVGVFHLRSPLGQYKLTFDKAREACANEAATMATYNQLSYAQKAKYHLCSAGW
LETGRVAYPTAFASQNCGSGVVGIVDYGPRPNKSEMWDVFCYRMKDVNCTCKVGYVGD
GFSCSGNLLQVLMSFPSLTNFLTEVLAYSNSSARGRAFLEHLTDLSIRGTLFVPQNSG
LGENETLSGRDIEHHLANVSMFFYNDLVNGTTLQTRLGSKLLITASQDPLQPTETRFV
DGRAILQWDIFASNGIIHVISRPLKAPPAPVTLTHTGLGAGIFFAIILVTGAVALAAY
SYFRINRRTIGFQHFESEEDINVAALGKQQPENISNPLYESTTSAPPEPSYDPFTDSE
ERQLEGNDPLRTL"

ORIGIN

Query Match 99.2%; Score 3234.2; DB 9; Length 8251;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3239; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy       7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5003 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5062

Qy      67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5063 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 5122

Qy     127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5123 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 5182

Qy     187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5183 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 5242

Qy     247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5243 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 5302

Qy     307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5303 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 5362

Qy     367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5363 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 5422

Qy     427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5423 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 5482

Qy     487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5483 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 5542

Qy     547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5543 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 5602

Qy     607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5603 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 5662

Qy     667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5663 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 5722

Qy     727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5723 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 5782

Qy     787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5783 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 5842

Qy     847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5843 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 5902

Qy     907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5903 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 5962

Qy     967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5963 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 6022

Qy    1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6023 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 6082

Qy    1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6083 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 6142

Qy    1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6143 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 6202

Qy    1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6203 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 6262

Qy    1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6263 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 6322

Qy    1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6323 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 6382

Qy    1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6383 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 6442

Qy    1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6443 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 6502

Qy    1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
```

```
Db       6503 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 6562

Qy       1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626

Db       6563 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 6622

Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686

Db       6623 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 6682

Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746

Db       6683 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 6742

Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806

Db       6743 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 6802

Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866

Db       6803 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 6862

Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926

Db       6863 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 6922

Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986

Db       6923 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 6982

Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046

Db       6983 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 7042

Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106

Db       7043 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 7102

Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166

Db       7103 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 7162

Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226

Db       7163 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 7222

Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286

Db       7223 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 7282

Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346

Db       7283 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 7342

Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406

Db       7343 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 7402

Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466

Db       7403 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 7462

Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526

Db       7463 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 7522

Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586

Db       7523 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 7582

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646

Db       7583 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 7642

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706

Db       7643 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 7702

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766

Db       7703 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 7762

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826

Db       7763 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 7822

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886

Db       7823 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 7882

Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946

Db       7883 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 7942

Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006

Db       7943 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 8002

Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066

Db       8003 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 8062

Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126

Db       8063 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 8122

Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186

Db       8123 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 8182

Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246

Db       8183 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 8242

Qy       3247 AAAAAAAA 3254

Db       8243 AGTCAACA 8250
```

```
RESULT 5
AK074051
LOCUS       AK074051               5604 bp    mRNA    linear   PRI 15-FEB-2002
DEFINITION  Homo sapiens mRNA for FLJ00122 protein.
ACCESSION   AK074051
VERSION     AK074051.1  GI:18676449
KEYWORDS    fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
  TITLE     The nucleotide sequence of a long cDNA clone isolated from human
            spleen
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 5604)
  AUTHORS   Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JAN-2002) Takahiro Nagase, Kazusa DNA Research
            Institute, Department of Human Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913,
            Fax:81-438-52-3914)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert and 5'- &
```

```
          3'-end one pass sequencing: Research Association for Biotechnology;
          cDNA library construction and clone selection: Kazusa DNA Research
          Institute.
FEATURES             Location/Qualifiers
     source          1. .5604
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="FLJ00122"
                     /tissue_type="spleen"
                     /dev_stage="adult"
                     /note="vector:pBluescriptII SK plus"
     gene            1. .5604
                     /gene="FLJ00122"
     CDS             <1. .5213
                     /gene="FLJ00122"
                     /note="Start codon is not identified."
                     /codon_start=3
                     /evidence=not_experimental
                     /product="FLJ00122 protein"
                     /protein_id="BAB84877.1"
                     /db_xref="GI:18676450"
                     /translation="DGSAGRLCDKQTSACGPYVQFCHIHATCEYSNGTASCICKAGYE
                     GDGTLCSEMDPCTGLTPGGCSRNAECIKTGTGTHTCVCQQGWTGNGRDCSEINNCLLP
                     SAGGCHDNASCLYVGPGQNECECKKGFRGNGIDCEPITSCLEQTGKCHPLASCQSTSS
                     GVWSCVCQEGYEGDGFLCYGNAAVELSFLSEAAIFNRWINNASLQPTLSATSNLTVLV
                     PSQQATEDMDQDEKSFWLSQSNIPALIKYHMLLGTYRVADLQTLSSSDMLATSLQGNF
                     LHLAKVDGNITIEGASIVDGDNAATNGVIHIINKVLVPQRRLTGSLPNLLMRLEQMPD
                     YSIFRGYIIQYNLANAIEAADAYTVFAPNNNAIENYIREKKVLSLEEDVLRYHVVLEE
                     KLLKNDLHNGMHRETMLGFSYFLSFFLHNDQLYVNEAPINYTNVATDKGVIHGLGKVL
                     EIQKNRCDNNDTTIIRGRCRTCSSELTCPFGTKSLGNEKRRCIYTSYFMGRRTLFIGC
                     QPKCVRTVITRECCAGFFGPQCQPCPGNAQNVCFGNGICLDGVNGTGVCECGEGFSGT
                     ACETCTEGKYGIHCDQACSCVHGRCNQGPLGDGSCDCDVGWRGVHCDNATTEDNCNGT
                     CHTSANCLTNSDGTASCKCAAGFQGNGTICTAINACEISNGGCSAKADCKRTTPGRRV
                     CTCKAGYTGDGIVCLEINPCLENHGGCDKNAECTQTGPNQAACNCLPAYTGDGKVCTL
                     INVCLTKNGGCSEFAICNHTGQVERTCTCKPNYIGDGFTCRGSIYQELPKNPKTSQYF
                     FQLQEHFVKDLVGPGPFTVFAPLSAAFDEEARVKDWDKYGLMPQVLRYHVVACHQLLL
                     ENLKLISNATSLQGEPIVISVSQSTVYINNKAKIISSDIISTNGIVHIIDKLLSPKNL
                     LITPKDNSGRILQNLTTLATNNGYIKFSNLIQDSGLLSVITDPIHTPVTLFWPTDQAL
                     HALPAEQQDFLFNQDNKDKLKEYLKFHVIRDAKVLAVDLPTSTAWKTLQGSELSVKCG
                     AGRDIGDLFLNGQTCRIVQRELLFDLGVAYGIDCLLIDPTLGGRCDTFTTFDASGECG
                     SCVNTPSCPRWSKPKGVKQKCLYNLPFKRNLEGCRERCSLVIQIPRCCKGYFGRDCQA
                     CPGGPDAPCNNRGVCLDQYSATGECKCNTGFNGTACEMCWPGRFGPDCLPCGCSDHGQ
                     CDDGITGSGQCLCETGWTGTSCDTQAVLPAVCTPPCSAHATCKENNTCECNLDYEGDG
                     ITCTVVDFCKQDNGGCAKVARCSQKGTKVSCSCQKGYKGDGHSCTEIDPCADGLNGGC
                     HEHATCKMTGPGKHKCECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQ
                     DTTVGVFHLRSPLGQYKLTFDKAREACANEAATMATYNQLSYAQKAKYHLCSAGWLET
                     GRVAYPTAFASQNCGSGVVGIVDYGPRPNKSEMWDVFCYRMKDVNCTCKVGYVGDGFS
                     CSGNLLQVLMSFPSLTNFLTEVLAYSNSSARGRAFLEHLTDLSIRGTLFVPQNSGLGE
                     NETLSGRDIEHHLANVSMFFYNDLVNGTTLQTRLGSKLLITASQDPLQPTETRFVDGR
                     AILQWDIFASNGIIHVISRPLKAPPAPVTLTHTGLGAGIFFAIILVTGAVALAAYSYF
                     RINRRTIGFQHFESEEDINVAALGKQQPENISNPLYESTTSAPPEPSYDPFTDSEERQ
                     LEGNDPLRTL"
ORIGIN

  Query Match           99.2%;  Score 3232.6;  DB 9;  Length 5604;
  Best Local Similarity 99.7%;  Pred. No. 0;
  Matches 3238;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2356 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 2415

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2416 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 2475

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2476 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 2535

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       2536 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 2595

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2596 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 2655

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2656 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 2715

Qy        367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2716 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 2775

Qy        427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2776 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 2835

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2836 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 2895

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2896 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 2955

Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2956 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 3015

Qy        667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3016 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 3075

Qy        727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3076 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 3135

Qy        787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3136 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 3195

Qy        847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3196 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 3255

Qy        907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3256 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 3315

Qy        967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3316 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 3375

Qy       1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3376 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 3435

Qy       1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3436 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 3495

Qy       1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3496 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 3555

Qy       1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3556 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 3615

Qy       1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       3616 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCACCT 3675
```

```
Qy       1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
Db       3676 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 3735

Qy       1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
Db       3736 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 3795

Qy       1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
Db       3796 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 3855

Qy       1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
Db       3856 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 3915

Qy       1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
Db       3916 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 3975

Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
Db       3976 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 4035

Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
Db       4036 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 4095

Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
Db       4096 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 4155

Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
Db       4156 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 4215

Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
Db       4216 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 4275

Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
Db       4276 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 4335

Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
Db       4336 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 4395

Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
Db       4396 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 4455

Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
Db       4456 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 4515

Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
Db       4516 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 4575

Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
Db       4576 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 4635

Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
Db       4636 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 4695

Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
Db       4696 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 4755

Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
Db       4756 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 4815

Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
Db       4816 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 4875

Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
Db       4876 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 4935

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
Db       4936 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 4995

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
Db       4996 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 5055

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
Db       5056 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 5115

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
Db       5116 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 5175

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
Db       5176 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 5235

Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
Db       5236 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 5295

Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
Db       5296 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 5355

Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
Db       5356 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 5415

Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126
Db       5416 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 5475

Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186
Db       5476 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 5535

Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246
Db       5536 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 5595

Qy       3247 AAAAAAAA 3254
Db       5596 AGTCAACA 5603
```

```
RESULT 6
AK024503
LOCUS       AK024503                4575 bp    mRNA    linear   PRI 29-SEP-2000
DEFINITION  Homo sapiens mRNA for FLJ00112 protein, partial cds.
ACCESSION   AK024503
VERSION     AK024503.1  GI:10440525
KEYWORDS    fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4575)
```

```
  AUTHORS   Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
  TITLE     The nucleotide sequence of a long cDNA clone isolated from human
            spleen
  JOURNAL   Published Only in DataBase (2000)
REFERENCE   2  (bases 1 to 4575)
  AUTHORS   Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,
            Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913,
            Fax:81-438-52-3914)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Kazusa DNA Research Institute.
FEATURES             Location/Qualifiers
     source          1. .4575
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="as00112"
                     /tissue_type="spleen"
                     /dev_stage="adult"
                     /note="vector:pBluescriptII SK plus"
     gene            1. .4575
                     /gene="FLJ00112"
     CDS             <606. .4184
                     /gene="FLJ00112"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /product="FLJ00112 protein"
                     /protein_id="BAB15793.1"
                     /db_xref="GI:10440526"
                     /translation="WHLFGWSDGTGVCECGEGFSGTACETCTEGKYGIHCDQACSCVH
                     GRCNQGPLGDGSCDCDVGWRGVHCDNATTEDNCNGTCHTSANCLTNSDGTASCKCAAG
                     FQGNGTICTAINACEISNGGCSAKADCKRTTPGRRVCTCKAGYTGDGIVCLEINPCLE
                     NHGGCDKNAECTQTGPNQAACNCLPAYTGDGKVCTLINVCLTKNGGCSEFAICNHTGQ
                     VERTCTCKPNYIGDGFTCRGSIYQELPKNPKTSQYFFQLQEHFVKDLVGPGPFTVFAP
                     LSAAFDEEARVKDWDKYGLMPQVLRYHVVACHQLLLENLKLISNATSLQGEPIVISVS
                     QSTVYINNKAKIISSDIISTNGIVHIIDKLLSPKNLLITPKDNSGRILQNLTTLATNN
                     GYIKFSNLIQDSGLLSVITDPIHTPVTLFWPTDQALHALPAEQQDFLFNQDNKDKLKE
                     YLKFHVIRDAKVLAVDLPTSTAWKTLQGSELSVKCGAGRDIGDLFLNGQTCRIVQREL
                     LFDLGVAYGIDCLLIDPTLGGRCDTFTTFDASGECGSCVNTPSCPRWSKPKGVKQKCL
                     YNLPFKRNLEGCRERCSLVIQIPRCCKGYFGRDCQACPGGPDAPCNNRGVCLDQYSAT
                     GECKCNTGFNGTACEMCWPGRFGPDCLPCGCSDHGQCDDGITGSGQCLCETGWTGPSC
                     DTQAVLPAVCTPPCSAHATCKENNTCECNLDYEGDGITCTVVDFCKQDNGGCAKVARC
                     SQKGTKVSCSCQKGYKGDGHSCTEIDPCADGLNGGCHEHATCKMTGPGKHKCECKSHY
                     VGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHLRSPLGQYKLTFDK
                     AREACANEAATMATYNQLSYAQKAKYHLCSAGWLETGRVAYPTAFASQNCGSGVVGIV
                     DYGPRPNKSEMWDVFCYRMKDVNCTCKVGYVGDGFSCSGNLLQVLMSFPSLTNFLTEV
                     LAYSNSSARGRAFLEHLTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLANVSMFFYN
                     DLVNGTTLQTRLGSKLLITASQDPLQPTETRFVDGRAILQWDIFASNGIIHVISRPLK
                     APPAPVTLTHTGLGAGIFFAIILVTGAVALAAYSYFRINRRTIGFQHFESEEDINVAA
                     LGKQQPENISNPLYESTTSAPPEPSYDPFTDSEERQLEGNDPLRTL"
ORIGIN

  Query Match           99.1%;  Score 3231;  DB 9;  Length 4575;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 3237;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1327 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCACT 1386

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 1446

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1447 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 1506

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1507 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 1566

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1567 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 1626

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1627 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 1686

Qy        367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1687 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 1746

Qy        427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 1806

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1807 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 1866

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1867 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 1926

Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1927 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 1986

Qy        667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1987 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 2046

Qy        727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2047 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 2106

Qy        787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2107 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 2166

Qy        847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2167 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 2226

Qy        907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2227 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 2286

Qy        967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2287 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 2346

Qy       1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2347 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 2406

Qy       1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2407 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 2466

Qy       1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2467 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 2526

Qy       1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2527 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 2586
```

```
Qy       1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
Db       2587 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 2646

Qy       1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
Db       2647 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 2706

Qy       1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
Db       2707 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 2766

Qy       1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
Db       2767 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 2826

Qy       1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
Db       2827 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 2886

Qy       1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
Db       2887 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 2946

Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
Db       2947 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 3006

Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
Db       3007 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 3066

Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
Db       3067 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 3126

Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
Db       3127 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 3186

Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
Db       3187 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 3246

Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
Db       3247 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 3306

Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
Db       3307 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 3366

Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
Db       3367 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 3426

Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
Db       3427 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 3486

Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
Db       3487 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 3546

Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
Db       3547 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 3606

Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
Db       3607 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 3666

Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
Db       3667 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 3726

Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
Db       3727 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 3786

Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
Db       3787 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 3846

Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
Db       3847 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 3906

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
Db       3907 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 3966

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
Db       3967 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 4026

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
Db       4027 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 4086

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
Db       4087 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 4146

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
Db       4147 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 4206

Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
Db       4207 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 4266

Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
Db       4267 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 4326

Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
Db       4327 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 4386

Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126
Db       4387 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 4446

Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186
Db       4447 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 4506

Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246
Db       4507 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 4566

Qy       3247 AAAAAAAA 3254
Db       4567 AGTCAACA 4574
```

```
RESULT 7
AB052958
LOCUS       AB052958               8214 bp    mRNA    linear   PRI 11-SEP-2002
DEFINITION  Homo sapiens mRNA for FELE-2, complete cds.
ACCESSION   AB052958
VERSION     AB052958.1  GI:22779440
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Adachi,H. and Tsujimoto,M.
TITLE FEEL-1, a novel scavenger receptor with in vitro bacteria-binding and angiogenesis-modulating activities
JOURNAL J. Biol. Chem. 277 (37), 34264-34270 (2002)
MEDLINE 22206516
PUBMED 12077138
REFERENCE 2 (bases 1 to 8214)
AUTHORS Adachi,H.
TITLE Direct Submission
JOURNAL Submitted (25-DEC-2000) Hideki Adachi, RIKEN, Cellular Biochemistry Laboratory; Hirosawa 2-1, Wako-shi, Saitama 351-0198, Japan (E-mail:adachih@postman.riken.go.jp, Tel:81-48-467-9372, Fax:81-48-462-4670)

FEATURES Location/Qualifiers
source 1. .8214
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1. .8214
/gene="FELE-2"
CDS 183. .7838
/gene="FELE-2"
/note="fasciclin, EGF and laminin type EGF containing scavenger receptor-2"
/codon_start=1
/product="FELE-2"
/protein_id="BAC15608.1"
/db_xref="GI:22779441"
/translation="MMLQHLVIFCLGLVVQNFCSPAETTGQARRCDRKSLLTIRTECR
SCALNLGVKCPDGYTMITSGSVGVRDCRYTFEVRTYSLSLPGCRHICRKDYLQPRCCP
GRWGPDCIECPGGAGSPCNGRGSCAEGMEGNGTCSCQEGFGGTACETCADDNLFGPSC
SSVCNCVHGVCNSGLDGDGTCECYSAYTGPKCDKPIPECAALLCPENSRCSPSTEDEN
KLECKCLPNYRGDGKYCDPINPCLRKICHPHAHCTYLGPNRHSCTCQEGYRGDGQVCL
PVDPCQINFGNCPTKSTVCKYDGPGQSHCECKEHYQNFVPGVGCSMTDICKSDNPCHR
NANCTTVAPGRTECICQKGYVGDGLTCYGNIMERLRELNTEPRGKWQGRLTSFISLLD
KAYAWPLSKLGPFTVLLPTDKGLKGFNVNELLVDNKAAQYFVKLHIIAGQMNIEYMNN
TDMFYTLTGKSGEIFNSDKDNQIKLKLHGGKKKVKIIQGDIIASNGLLHILDRAMDKL
EHTFESNNEQTIMTMLQPRYSKFRSLLEETNLGHALDEDGVGGPYTIFVPNNEALNNM
KDGTLDYLLSPEGSRKLLELVRYHIVPFTQLEVATLISTPHIRSMANQLIQFNTTDNG
QILANDVAMEEIEITAKNGRIYTLTGVLIPPSIVPILPHRCDETKREMKLGTCVSCSL
VYWSRCPANSEPTALFTHRCVYSGRFGSLKSGCARYCNATVKIPKCCKGFYGPDCNQC
PGGFSNPCSGNGQCADSLGGNGTCICEEGFQGSQCQFCSDPNKYGPRCNKKCLCVHGT
CNNRIDSDGACLTGTCRDGSAGRLCDKQTSACGPYVQFCHIHATCEYSNGTASCICKA
GYEGDGTLCSEMDPCTGLTPGGCSRNAECIKTGTGTHTCVCQQGWTGNGRDCSEINNC
LLPSAGGCHDNASCLYVGPGQNECECKKGFRGNGIDCEPITSCLEQTGKCHPLASCQS
TSSGVWSCVCQEGYEGDGFLCYGNAAVELSFLSEAAIFNRWINNASLQPTLSATSNLT
VLVPSQQATEDMDQDEKSFWLSQSNIPALIKYHMLLGTYRVADLQTLSSSDMLATSLQ
GNFLHLAKVDGNITIEGASIVDGDNAATNGVIHIINKVLVPQRRLTGSLPNLLMRLEQ
MPDYSIFRGYIIQYNLANAIEAADAYTVFAPNNNAIENYIREKKVLSLEEDVLRYHVV
LEEKLLKNDLHNGMHRETMLGFSYFLSFFLHNDQLYVNEAPINYTNVATDKGVIHGLG
KVLEIQKNRCDNNDTTIIRGRCRTCSSELTCPFGTKSLGNEKRRCIYTSYFMGRRTLF
IGCQPKCVRTVITRECCAGFFGPQCQPCPGNAQNVCFGNGICLDGVNGTGVCECGEGF
SGTACETCTEGKYGIHCDQACSCVHGRCNQGPLGDGSCDCDVGWRGVHCDNATTEDNC
NGTCHTSANCLTNSDGTASCKCAAGFQGNGTICTAINACEISNGGCSAKADCKRTTPG
RRVCTCKAGYTGDGIVCLEINPCLENHGGCDKNAECTQTGPNQAACNCLPAYTGDGKV
CTLINVCLTKNGGCSEFAICNHTGQVERTCTCKPNYIGDGFTCRGSIYQELPKNPKTS
QYFFQLQEHFVKDLVGPGPFTVFAPLSAAFDEEARVKDWDKYGLMPQVLRYHVVACHQ
LLLENLKLISNATSLQGEPIVISVSQSTVYINNKAKIISSDIISTNGIVHIIDKLLSP
KNLLITPKDNSGRILQNLTTLATNNGYIKFSNLIQDSGLLSVITDPIHTPVTLFWPTD
QALHALPAEQQDFLFNQDNKDKLKEYLKFHVIRDAKVLAVDLPTSTAWKTLQGSELSV
KCGAGRDIGDLFLNGQTCRIVQRELLFDLGVAYGIDCLLIDPTLGGRCDTFTTFDASG
ECGSCVNTPSCPRWSKPKGVKQKCLYNLPFKRNLEGCRERCSLVIQIPRCCKGYFGRD
CQACPGGPDAPCNNRGVCLDQYSATGECKCNTGFNGTACEMCWPGRFGPDCLPCGCSD
HGQCDDGITGSGQCLCETGWTGPSCDTQAVLPAVCTPPCSAHATCKENNTCECNLDYE
GDGITCTVVDFCKQDNGGCAKVARCSQKGTKVSCSCQKGYKGDGHSCTEIDPCADGLN
GGCHEHATCKMTGPGKHKCECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDL
HFQDTTVGVFHLRSPLGQYKLTFDKAREACANEAATMATYNQLSYAQKAKYHLCSAGW
LETGRVAYPTAFASQNCGSGVVGIVDYGPRPNKSEMWDVFCYRMKDVNCTCKVGYVGD
GFSCSGNLLQVLMSFPSLTNFLTEVLAYSNSSARGRAFLEHLTDLSIRGTLFVPQNSG
LGENETLSGRDIEHHLANVSMFFYNDLVNGTTLQTRLGSKLLITASQDPLQPTETRFV
DGRAILQWDIFASNGIIHVISRPLKAPPAPVTLTHTGLGAGIFFAIILVTGAVALAAY
SYFRINRRTIGFQHFESEEDINVAALGKQQPENISNPLYESTTSAPPEPSYDPFTDSE
ERQLEGNDPLRTL"

ORIGIN

Query Match 99.1%; Score 3229.8; DB 9; Length 8214;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3231; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4981 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5040

Qy      67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5041 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 5100

Qy     127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5101 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 5160

Qy     187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5161 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 5220

Qy     247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5221 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 5280

Qy     307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5281 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 5340

Qy     367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5341 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 5400

Qy     427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5401 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 5460

Qy     487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5461 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 5520

Qy     547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5521 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 5580

Qy     607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5581 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 5640

Qy     667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5641 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 5700

Qy     727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5701 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 5760

Qy     787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5761 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 5820

Qy     847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5821 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 5880

Qy     907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5881 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 5940
```

```
Qy        967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5941 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 6000

Qy       1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6001 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 6060

Qy       1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6061 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 6120

Qy       1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6121 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 6180

Qy       1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6181 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 6240

Qy       1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6241 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 6300

Qy       1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6301 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 6360

Qy       1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6361 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 6420

Qy       1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6421 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 6480

Qy       1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6481 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 6540

Qy       1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6541 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 6600

Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6601 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 6660

Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6661 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 6720

Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6721 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 6780

Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6781 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 6840

Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6841 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 6900

Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6901 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 6960

Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6961 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 7020

Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7021 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 7080

Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7081 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 7140

Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7141 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 7200

Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7201 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 7260

Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7261 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 7320

Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7321 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 7380

Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7381 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 7440

Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7441 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 7500

Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7501 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 7560

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7561 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 7620

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7621 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 7680

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7681 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 7740

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7741 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 7800

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7801 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 7860

Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7861 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 7920

Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7921 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 7980

Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7981 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 8040

Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8041 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 8100

Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
```

```
Db       8101 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 8160

Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8161 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 8214


RESULT 8
AY227444
LOCUS       AY227444                4492 bp    mRNA    linear   PRI 23-APR-2003
DEFINITION  Homo sapiens hyaluronan receptor for endocytosis precursor (HARE)
            mRNA, partial cds.
ACCESSION   AY227444
VERSION     AY227444.1  GI:28848629
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4492)
  AUTHORS   Zhou,B., McGary,C.T., Weigel,J.A., Saxena,A. and Weigel,P.H.
  TITLE     Purification and molecular identification of the human hyaluronan
            receptor for endocytosis
  JOURNAL   Glycobiology 13 (5), 339-349 (2003)
  MEDLINE   22588541
   PUBMED   12626425
REFERENCE   2  (bases 1 to 4492)
  AUTHORS   Weigel,P.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JAN-2003) Biochemistry and Molecular Biology,
            University of Oklahoma Health Sciences Center, BMSB Room 860,
            Oklahoma City, OK 73104, USA
FEATURES             Location/Qualifiers
     source          1. .4492
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            <1. .4492
                     /gene="HARE"
     CDS             <1. .4251
                     /gene="HARE"
                     /codon_start=1
                     /product="hyaluronan receptor for endocytosis precursor"
                     /protein_id="AAO39681.1"
                     /db_xref="GI:28848630"
                     /translation="SLPNLLMRLEQMPDYSIFRGYIIQYNLANAIEAADAYTVFAPNN
                     NAIENYIREKKVLSLEEDVLRYHVVLEEKLLKNDLHNGMHRETMLGFSYFLSFFLHND
                     QLYVNEAPINYTNVATDKGVIHGLGKVLEIQKNRCDNNDTTIIRGRCRTCSSELTCPF
                     GTKSLGNEKRRCIYTSYFMGRRTLFIGCQPKCVRTVITRECCAGFFGPQCQPCPGNAQ
                     NVCFGNGICLDGVNGTGVCECGEGFSGTACETCTEGKYGIHCDQACSCVHGRCNQGPL
                     GDGSCDCDVGWRGVHCDNATTEDNCNGTCHTSANCLTNSDGTASCKCAAGFQGNGTIC
                     TAINACEISNGGCSAKADCKRTTPGRRVCTCKAGYTGDGIVCLEINPCLENHGGCDKN
                     AECTQTGPNQAACNCLPAYTGDGKVCTLIYVCLTKNGGCSEFAICNHTGQVERTCTCK
                     PNYIGDGFTCRGSIYQELPKNPKTSQYFFQLQEHFVKDLVGPGPFTVFAPLSAAFDEE
                     ARVKDWDKYGLMPQVLRYHVVACHQLLLENLKLISNATSLQGEPIVISVSQSTVYINN
                     KAKIISSDIISTNGIVHIIDKLLSPKNLLITPKDNSGRILQNLTTLATNNGYIKFSNL
                     IQDSGLLSVITDPIHTPVTLFWPTDQALHALPAEQQDFLFNQDNKDKLKEYLKFHVIR
                     DAKVLAVDLPTSTAWKTLQGSELSVKCGAGRDIGDLFLNGQTCRIVQRELLFDLGVAY
                     GIDCLLIDPTLGGRCDTFTTFDASGECGSCVNTPSCPRWSKPKGVKQKCLYNLPFKRN
                     LEGCRERCSLVIQIPRCCKGYFGRDCQACPGGPDAPCNNRGVCLDQYSATGECKCNTG
                     FNGTACEMCWPGRFGPDCLPCGCSDHGQCDDGITGSGQCLCETGWTGPSCDTQAVLPA
                     VCTPPCSAHATCKENNTCECNLDYEGDGITCTVVDFCKQDNGGCAKVARCSQKGTKVS
                     CSCQKGYKGDGHSCTEIDPCADGLNGGCHEHATCKMTGPGKHKCECKSHYVGDGLNCE
                     PEQLPIDRCLQDNGQCHADAKCVDLHFQDTTVGVFHLRSPLGQYKLTFDKAREACANE
                     AATMATYNQLSYAQKAKYHLCSAGWLETGRVAYPTAFASQNCGSGVVGIVDYGPRPNK
                     SEMWDVFCYRMKDVNCTCKVGYVGDGFSCSGNLLQVLMSFPSLTNFLTEVLAYSNSSA
                     RGRAFLEHLTDLSIRGTLFVPQNSGLGENETLSGRDIEHHLANVSMFFYNDLVNGTTL
                     QTRLGSKLLITASQDPLQPTETRFVDGRAILQWDIFASNGIIHVISRPLKAPPAPVTL
                     THTGLGAGIFFAIILVTGAVALAAYSYFRINRRTIGFQHFESEEDINVAALGKQQPEN
                     ISNPLYESTTSAPPEPSYDPFTDSEERQLEGNDPLRTL"
     mat_peptide     1. .4248
                     /gene="HARE"
                     /product="190 kDa hyaluronan receptor for endocytosis"
                     /function="mediates endocytosis of hyaluronan or
                     chondroitin sulfate via the clathrin coated pit pathway"
                     /note="generated by proteolytic processing of a larger
                     protein; probably a product of the HARE/Stabilin-2 gene;
                     190 kDa protein purified from spleen"
ORIGIN

  Query Match           94.9%;  Score 3094.2;  DB 9;  Length 4492;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 3096;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1394 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 1453

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 1513

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 1573

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1574 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 1633

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1634 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 1693

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1694 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 1753

Qy        367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1754 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 1813

Qy        427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1814 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 1873

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1874 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 1933

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1934 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 1993

Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1994 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 2053

Qy        667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2054 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 2113

Qy        727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2114 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 2173

Qy        787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2174 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 2233

Qy        847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2234 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 2293

Qy        907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      2294 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 2353
Qy       967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
Db      2354 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 2413
Qy      1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
Db      2414 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 2473
Qy      1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
Db      2474 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 2533
Qy      1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
Db      2534 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 2593
Qy      1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
Db      2594 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 2653
Qy      1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
Db      2654 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 2713
Qy      1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
Db      2714 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 2773
Qy      1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
Db      2774 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 2833
Qy      1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
Db      2834 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 2893
Qy      1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
Db      2894 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 2953
Qy      1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
Db      2954 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 3013
Qy      1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
Db      3014 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 3073
Qy      1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
Db      3074 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 3133
Qy      1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
Db      3134 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 3193
Qy      1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
Db      3194 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 3253
Qy      1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
Db      3254 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 3313
Qy      1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
Db      3314 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 3373
Qy      1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
Db      3374 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 3433
```

```
Qy      2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
Db      3434 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 3493
Qy      2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
Db      3494 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 3553
Qy      2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
Db      3554 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 3613
Qy      2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
Db      3614 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 3673
Qy      2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
Db      3674 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 3733
Qy      2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
Db      3734 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 3793
Qy      2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
Db      3794 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 3853
Qy      2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
Db      3854 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 3913
Qy      2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
Db      3914 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 3973
Qy      2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
Db      3974 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 4033
Qy      2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
Db      4034 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 4093
Qy      2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
Db      4094 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 4153
Qy      2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
Db      4154 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 4213
Qy      2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
Db      4214 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 4273
Qy      2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
Db      4274 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 4333
Qy      2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
Db      4334 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 4393
Qy      3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
Db      4394 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 4453
Qy      3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACC 3105
Db      4454 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACC 4492
```

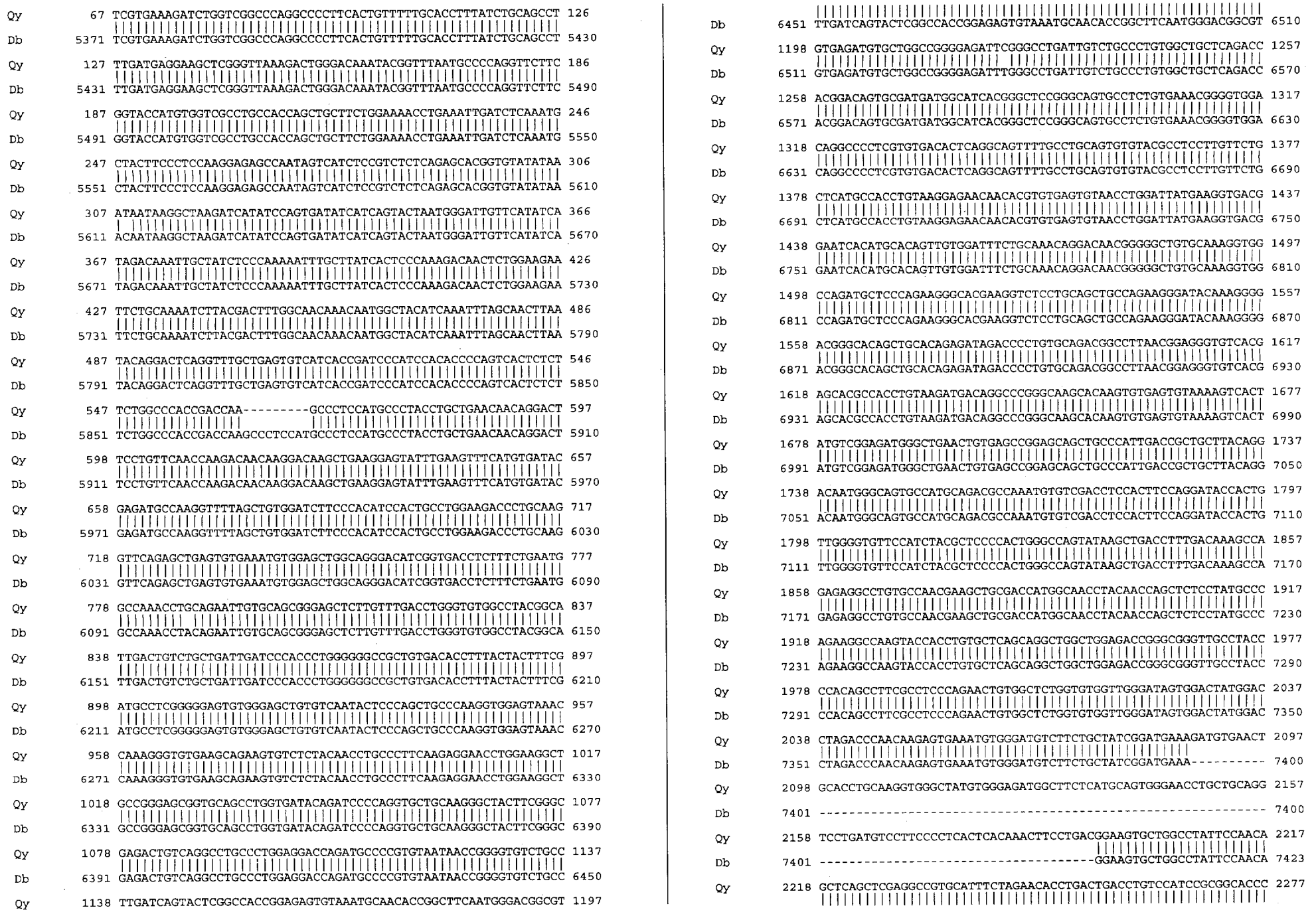
```
RESULT 9
HSM801377
LOCUS       HSM801377               3642 bp    mRNA    linear   PRI 18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp434E0321 (from clone DKFZp434E0321).
ACCESSION   AL133021
VERSION     AL133021.1  GI:6453577
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3642)
  AUTHORS   Blum,H., Bauersachs,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by LMU (Ludwig Maximilians University,
            Munich/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp434E0321) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .3642
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434E0321"
                     /tissue_type="testis"
                     /clone_lib="434 (synonym: htes3). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     gene            join(7. .2607,2609. .3217)
                     /gene="DKFZp434E0321"
     CDS             join(<7. .2607,2609. .3217)
                     /gene="DKFZp434E0321"
                     /note="strong similarity to KIAA0246
                     The frame shift was determined manually"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAB61358.1"
                     /db_xref="GI:6453578"
                     /db_xref="GOA:Q9UF98"
                     /db_xref="SPTREMBL:Q9UF98"
                     /translation="VGEAVGTASCKCAAGFQGNGTICTAINACEISNGGCSAKADCKR
                     TTPGRRVCTCKAGYTGDGIVCLEINPCLENHGGCDKNAECTQTGPNQAACNCLPAYTG
                     DGKVCTLINVCLTKNGGCSEFAICNHTGQVERTCTCKPNYIGDGFTCRGSIYQELPKN
                     PKTSQYFFQLQEHFVKDLVGPGPFTVFAPLSAAFDEEARVKDWDKYGLMPQVLRYHVV
                     ACHQLLLENLKLISNATSLQGEPIVISVSQSTVYINNKAKIISSDIISTNGIVHIIDK
                     LLSPKNLLITPKDNSGRILQNLTTLATNNGYIKFSNLIQDSGLLSVITDPIHTPVTLF
                     WPTDQALHALPAEQQDFLFNQDNKDKLKEYLKFHVIRDAKVLAVDLPTSTAWKTLQGS
                     ELSVKCGAGRDIGDLFLNGQTCRIVQRELLFDLGVAYGIDCLLIDPTLGGRCDTFTTF
                     DASGECGSCVNTPSCPRWSKPKGVKQKCLYNLPFKRNLEGCRERCSLVIQIPRCCKGY
                     FGRDCQACPGGPDAPCNNRGVCLDQYSATGECKCNTGFNGTACEMCWPGRFGPDCLPC
                     GCSDHGQCDDGITGSGQCLCETGWTGPSCDTQAVLPAVCTPPCSAHATCKENNTCECN
                     LDYEGDGITCTVVDFCKQDNGGCAKVARCSQKGTKVSCSCQKGYKGDGHSCTEIDPCA
                     DGLNGGCHEHATCKMTGPGKHKCECKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAK
                     CVDLHFQDTTVGVFHLRSPLGQYKLTFDKAREACANEAATMATYNQLSYAQKAKYHLC
                     SAGWLETGRVAYPTAFASQNCGSGVVGIVDYGPRPNKSEMWDVFCYRMKGSAGLFQQL
                     SSRPCISRTPDDLSIRGTLFVPQNSGLGENETLSGRDIEHHLANVSMFFYNDLVNGTT
                     LQTRLGSKLLITASQDPLQPTETRFVDGRAILQWDIFASNGIIHVISRPLKAPPAPVT
                     LTHTGLGAGIFFAIILVTGAVALAAYSYFRINRRTIGFQHFESEEDINVAALGKQQPE
                     NISNPLYESTTSAPPEPSYDPFTDSEERQLEGNDPLRTL"
     polyA_signal    3574. .3579
     polyA_site      3603
ORIGIN

  Query Match             92.6%;  Score 3017.8;  DB 9;  Length 3642;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches 3139;  Conservative    0;  Mismatches    8;  Indels  107;  Gaps    1;

Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 526

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 586

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 646

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 706

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 766

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 826

Qy        367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 886

Qy        427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 946

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 1006

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 1066

Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 1126

Qy        667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 1186

Qy        727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 1246

Qy        787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 1306

Qy        847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 1366

Qy        907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 1426

Qy        967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1486

Qy       1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1487 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1546
Qy       1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
Db       1547 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1606
Qy       1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
Db       1607 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1666
Qy       1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
Db       1667 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1726
Qy       1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
Db       1727 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1786
Qy       1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
Db       1787 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1846
Qy       1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
Db       1847 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1906
Qy       1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
Db       1907 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1966
Qy       1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
Db       1967 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 2026
Qy       1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
Db       2027 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 2086
Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
Db       2087 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 2146
Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
Db       2147 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 2206
Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
Db       2207 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 2266
Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
Db       2267 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 2326
Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
Db       2327 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 2386
Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
Db       2387 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 2446
Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
Db       2447 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2506
Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
Db       2507 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAA---------------------- 2547
Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
Db       2548 ------------------------------------------------------------ 2547

Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
Db       2548 ----------------------------GGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2579
Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
Db       2580 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2639
Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
Db       2640 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2699
Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
Db       2700 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2759
Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
Db       2760 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2819
Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
Db       2820 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2879
Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
Db       2880 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2939
Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
Db       2940 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2999
Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
Db       3000 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 3059
Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
Db       3060 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 3119
Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
Db       3120 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 3179
Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
Db       3180 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 3239
Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
Db       3240 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 3299
Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
Db       3300 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3359
Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
Db       3360 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3419
Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126
Db       3420 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3479
Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186
Db       3480 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 3539
Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246
Db       3540 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 3599
```

```
Qy       3247 AAAAAAAAAAAAAA 3260
              |    ||||||||||
Db       3600 AGTCAAAAAAAAAA 3613


RESULT 10
AX686192
LOCUS       AX686192                8444 bp    DNA     linear   PAT 29-MAR-2003
DEFINITION  Sequence 1 from Patent WO02059315.
ACCESSION   AX686192
VERSION     AX686192.1  GI:29372026
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A., Patturajan,M., Vernet,C.A., Casman,S.J.,
            Malyankar,U., Shenoy,S., Spytek,K.A., Gangolli,E., Miller,C.,
            Boldog,F., Li,L., Taupier,R.J., Kekuda,R., Smithson,G.,
            Zerhusen,B.D., Liu,X., Colman,S.D., Tchernev,V., Si,J., Edinger,S.,
            Stone,D., Sciore,P., Millet,I. and Rothenberg,M.
  TITLE     Human nucleic acids and polypeptides and methods of use thereof
  JOURNAL   Patent: WO 02059315-A 1 01-AUG-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .8444
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             92.0%;  Score 2999.8;  DB 6;  Length 8444;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches 3128;  Conservative    0;  Mismatches   18;  Indels  108;  Gaps    1;

Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5279 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5338

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5339 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 5398

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5399 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 5458

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5459 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 5518

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5519 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 5578

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5579 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 5638

Qy        367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5639 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 5698

Qy        427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5699 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 5758

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5759 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 5818

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5819 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 5878

Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5879 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 5938

Qy        667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5939 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 5998

Qy        727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5999 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 6058

Qy        787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6059 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 6118

Qy        847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6119 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 6178

Qy        907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6179 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 6238

Qy        967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6239 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 6298

Qy       1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6299 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 6358

Qy       1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6359 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 6418

Qy       1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6419 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 6478

Qy       1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       6479 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 6538

Qy       1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6539 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 6598

Qy       1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6599 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 6658

Qy       1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6659 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 6718

Qy       1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6719 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 6778

Qy       1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6779 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 6838

Qy       1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6839 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 6898

Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       6899 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 6958

Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6959 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 7018

Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7019 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 7078

Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7079 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 7138

Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7139 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 7198

Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7199 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 7258

Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7259 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 7318

Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
             ||||||||||||||||||||||||||||||||||||
Db       7319 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAA---------------------- 7358

Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
Db       7359 ------------------------------------------------------------ 7358

Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
                                      ||||||||||||||||||||||||||||||||||||
Db       7359 ----------------------------GGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 7390

Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7391 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 7450

Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7451 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 7510

Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7511 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 7570

Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
             || ||||||||||||||||||||||||||||||||||||||||||||||||    | | |
Db       7571 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGGTACAAAGTA 7630

Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7631 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 7690

Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7691 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 7750

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       7751 TGGGAGCAGGGATCTTCTTTTGCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 7810

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       7811 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTACCAGCATTTTGAGTCGGAAGAGG 7870

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7871 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 7930

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7931 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 7990

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7991 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 8050

Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8051 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 8110

Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8111 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 8170

Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8171 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 8230

Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8231 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 8290

Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       8291 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 8350

Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       8351 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 8410

Qy       3247 AAAAAAAAAAAAAA 3260
                |   ||||||||||
Db       8411 AGTCAAAAAAAAAA 8424
```

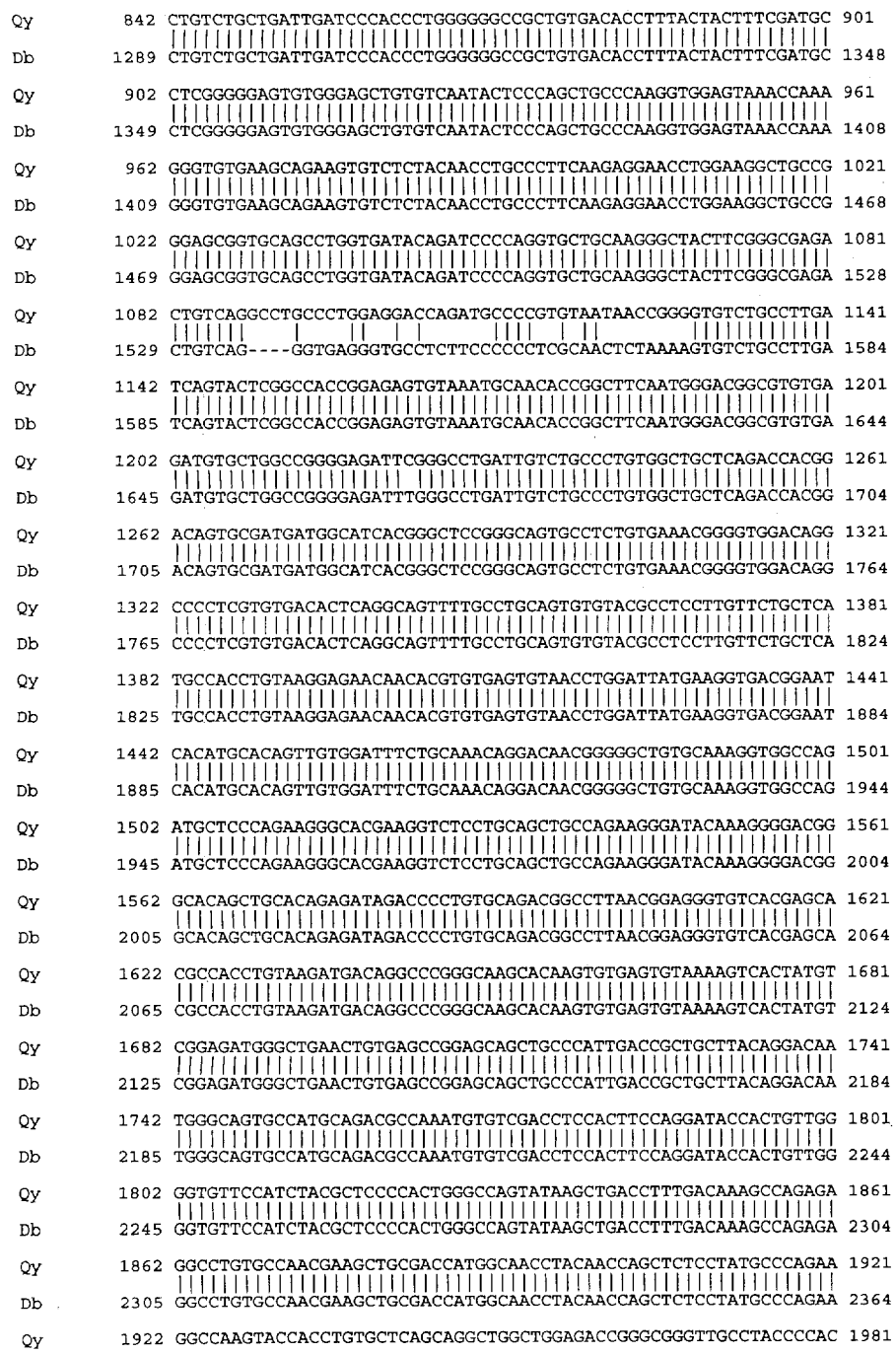
```
RESULT 11
AX686194
LOCUS       AX686194               8495 bp    DNA     linear   PAT 29-MAR-2003
DEFINITION  Sequence 3 from Patent WO02059315.
ACCESSION   AX686194
VERSION     AX686194.1  GI:29372027
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A., Patturajan,M., Vernet,C.A., Casman,S.J.,
            Malyankar,U., Shenoy,S., Spytek,K.A., Gangolli,E., Miller,C.,
            Boldog,F., Li,L., Taupier,R.J., Kekuda,R., Smithson,G.,
            Zerhusen,B.D., Liu,X., Colman,S.D., Tchernev,V., Si,J., Edinger,S.,
            Stone,D., Sciore,P., Millet,I. and Rothenberg,M.
  TITLE     Human nucleic acids and polypeptides and methods of use thereof
  JOURNAL   Patent: WO 02059315-A 3 01-AUG-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .8495
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           91.9%;  Score 2995.6;  DB 6;  Length 8495;
  Best Local Similarity   96.1%;  Pred. No. 0;
  Matches 3137;  Conservative    0;  Mismatches   10;  Indels  116;  Gaps    2;

Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5311 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5370
```



```
Db       7424 GCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCC 7483
Qy       2278 TCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCG 2337
Db       7484 TCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCG 7543
Qy       2338 AGCACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCC 2397
Db       7544 AGCACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCC 7603
Qy       2398 TGCAAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGA 2457
Db       7604 TGCAAACGAGGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGA 7663
Qy       2458 CGGAGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATG 2517
Db       7664 CGGAGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATG 7723
Qy       2518 GGATCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCC 2577
Db       7724 GGATCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCC 7783
Qy       2578 ACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCT 2637
Db       7784 ACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCT 7843
Qy       2638 TGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGT 2697
Db       7844 TGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGT 7903
Qy       2698 CGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACC 2757
Db       7904 CGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACC 7963
Qy       2758 CCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACT 2817
Db       7964 CCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACT 8023
Qy       2818 CTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGG 2877
Db       8024 CTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGG 8083
Qy       2878 AGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTT 2937
Db       8084 AGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTT 8143
Qy       2938 GCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 2997
Db       8144 GCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 8203
Qy       2998 ATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCC 3057
Db       8204 ATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCC 8263
Qy       3058 TCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTG 3117
Db       8264 TCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTG 8323
Qy       3118 CCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTG 3177
Db       8324 CCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTG 8383
Qy       3178 TATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAA 3237
Db       8384 TATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAA 8443
Qy       3238 CAGAAAAAAAAAAAAAAAAAAA 3260
Db       8444 CAGAAACAAAGTCAAAAAAAAAA 8466
```

RESULT 12
AX149461

```
LOCUS       AX149461               3625 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 17 from Patent WO0136638.
ACCESSION   AX149461
VERSION     AX149461.1  GI:14347936
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A., Lichenstein,H., Vernet,C. and Fernandes,E.
  TITLE     Polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0136638-A 17 25-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .3625
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             86.8%;  Score 2828.8;  DB 6;  Length 3625;
  Best Local Similarity   95.1%;  Pred. No. 0;
  Matches 2985;  Conservative    0;  Mismatches   43;  Indels  111;  Gaps    2;

Qy        122 AGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGT 181
Db        569 ACCCTCTGATCTTCTGTTACAGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGT 628
Qy        182 TCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTC 241
Db        629 TCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTC 688
Qy        242 AAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTA 301
Db        689 AAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTA 748
Qy        302 TATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCA 361
Db        749 TATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCA 808
Qy        362 TATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGG 421
Db        809 TATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGG 868
Qy        422 AAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAA 481
Db        869 AAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAA 928
Qy        482 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCAC 541
Db        929 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCAC 988
Qy        542 TCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCT 601
Db        989 TCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCT 1048
Qy        602 GTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 661
Db       1049 GTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 1108
Qy        662 TGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTC 721
Db       1109 TGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTC 1168
Qy        722 AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCA 781
Db       1169 AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCA 1228
Qy        782 AACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGA 841
Db       1229 AACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGA 1288
```

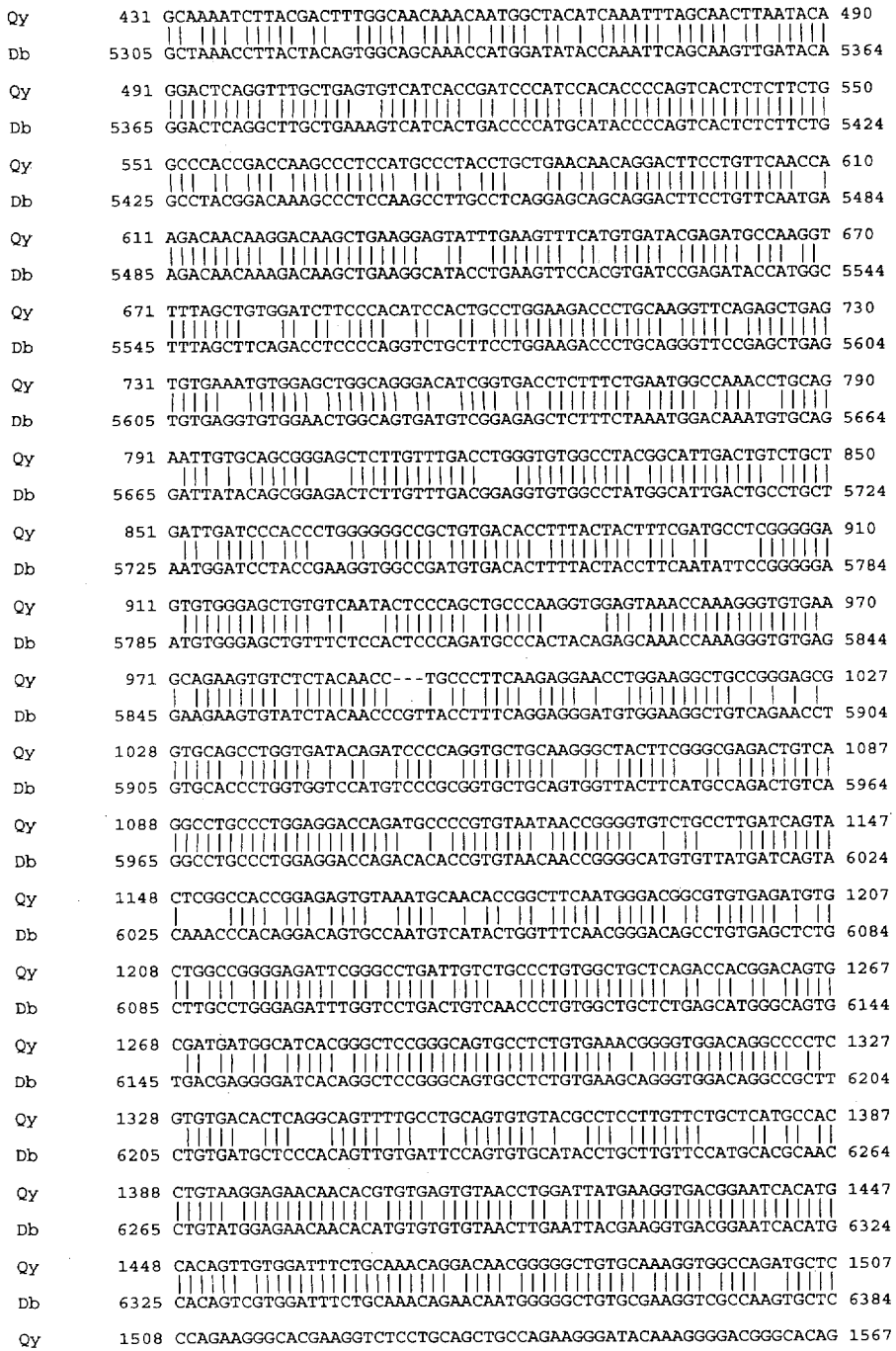

Qy 842 CTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGC 901
Db 1289 CTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGC 1348
Qy 902 CTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAA 961
Db 1349 CTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAA 1408
Qy 962 GGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCG 1021
Db 1409 GGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCG 1468
Qy 1022 GGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGA 1081
Db 1469 GGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGA 1528
Qy 1082 CTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGA 1141
Db 1529 CTGTCAG----GGTGAGGGTGCCTCTTCCCCCCTCGCAACTCTAAAAGTGTCTGCCTTGA 1584
Qy 1142 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA 1201
Db 1585 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA 1644
Qy 1202 GATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGG 1261
Db 1645 GATGTGCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGG 1704
Qy 1262 ACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGG 1321
Db 1705 ACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGG 1764
Qy 1322 CCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCA 1381
Db 1765 CCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCA 1824
Qy 1382 TGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAAT 1441
Db 1825 TGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAAT 1884
Qy 1442 CACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAG 1501
Db 1885 CACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAG 1944
Qy 1502 ATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGG 1561
Db 1945 ATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGG 2004
Qy 1562 GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCA 1621
Db 2005 GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCA 2064
Qy 1622 CGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGT 1681
Db 2065 CGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGT 2124
Qy 1682 CGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAA 1741
Db 2125 CGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAA 2184
Qy 1742 TGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGG 1801
Db 2185 TGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGG 2244
Qy 1802 GGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA 1861
Db 2245 GGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA 2304
Qy 1862 GGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAA 1921
Db 2305 GGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAA 2364
Qy 1922 GGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCAC 1981
Db 2365 GGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCAC 2424
Qy 1982 AGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAG 2041
Db 2425 AGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAG 2484
Qy 2042 ACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCAC 2101
Db 2485 ACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAA---------- 2530
Qy 2102 CTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCT 2161
Db 2531 ---------- 2530
Qy 2162 GATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTC 2221
Db 2531 ----------GGAAGTGCTGGCCTATTCCAACAGCTC 2557
Qy 2222 AGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTT 2281
Db 2558 AGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTT 2617
Qy 2282 TGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCA 2341
Db 2618 TGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCA 2677
Qy 2342 CCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCA 2401
Db 2678 CCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCA 2737
Qy 2402 AACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGA 2461
Db 2738 AACGAGGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGA 2797
Qy 2462 GACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGAT 2521
Db 2798 GACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGAT 2857
Qy 2522 CATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACAC 2581
Db 2858 CATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACAC 2917
Qy 2582 TGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGC 2641
Db 2918 TGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGC 2977
Qy 2642 TGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGA 2701
Db 2978 TGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGA 3037
Qy 2702 AGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTT 2761
Db 3038 AGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTT 3097
Qy 2762 GTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGA 2821
Db 3098 GTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGA 3157
Qy 2822 AGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGAT 2881
Db 3158 AGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGAT 3217
Qy 2882 GCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCT 2941
Db 3218 GCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCT 3277
Qy 2942 TTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCT 3001
Db 3278 TTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCT 3337
Qy 3002 GGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCT 3061

```
Db       3338 GGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCT 3397

Qy       3062 CTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCT 3121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3398 CTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCT 3457

Qy       3122 ACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATC 3181
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       3458 ACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATC 3517

Qy       3182 CCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGA 3241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3518 CCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGA 3577

Qy       3242 AAAAAAAAAAAAAAAAAAA 3260
              || |||   ||||||||||
Db       3578 AACAAAGTCAAAAAAAAAA 3596
```

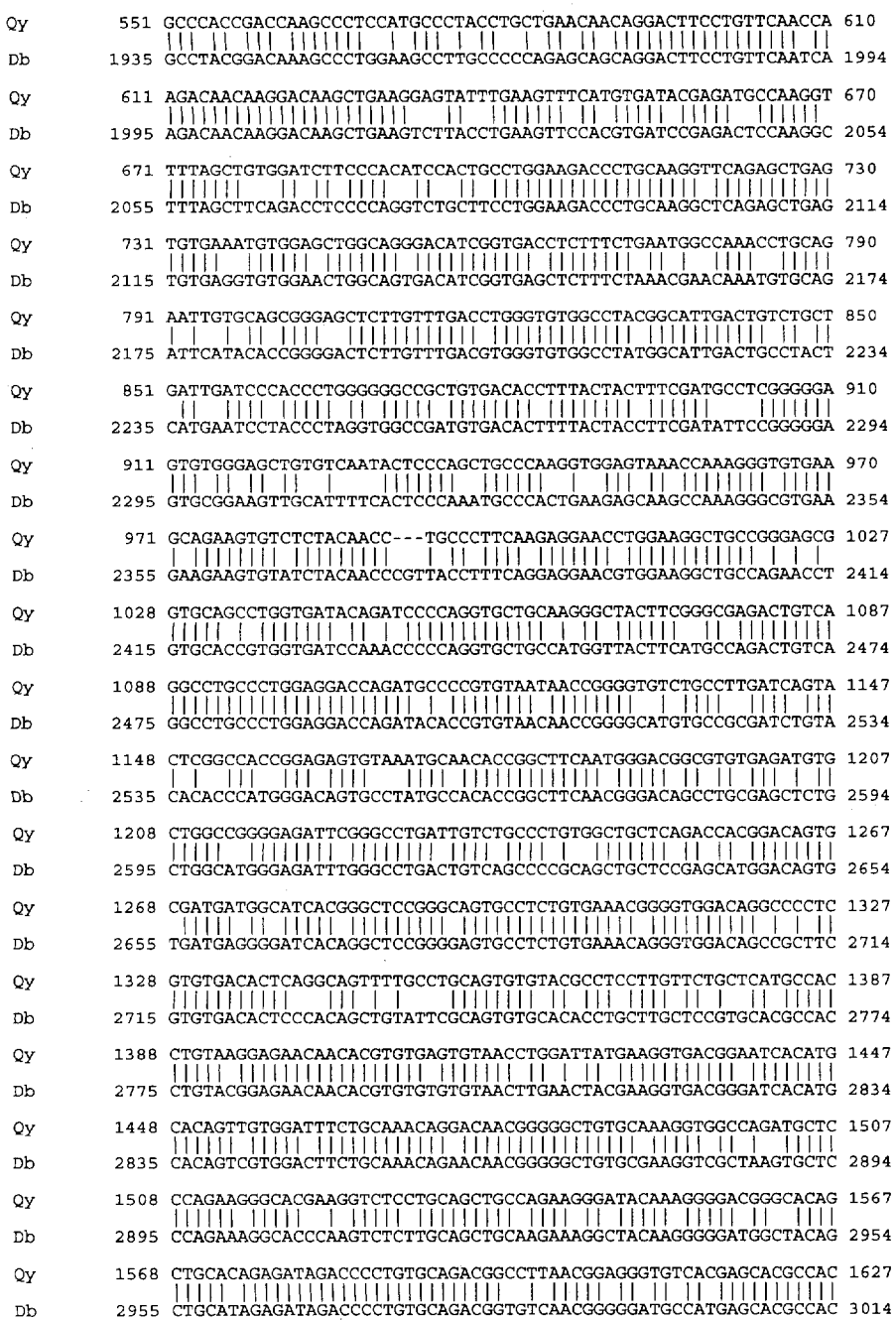
RESULT 13
AF364951

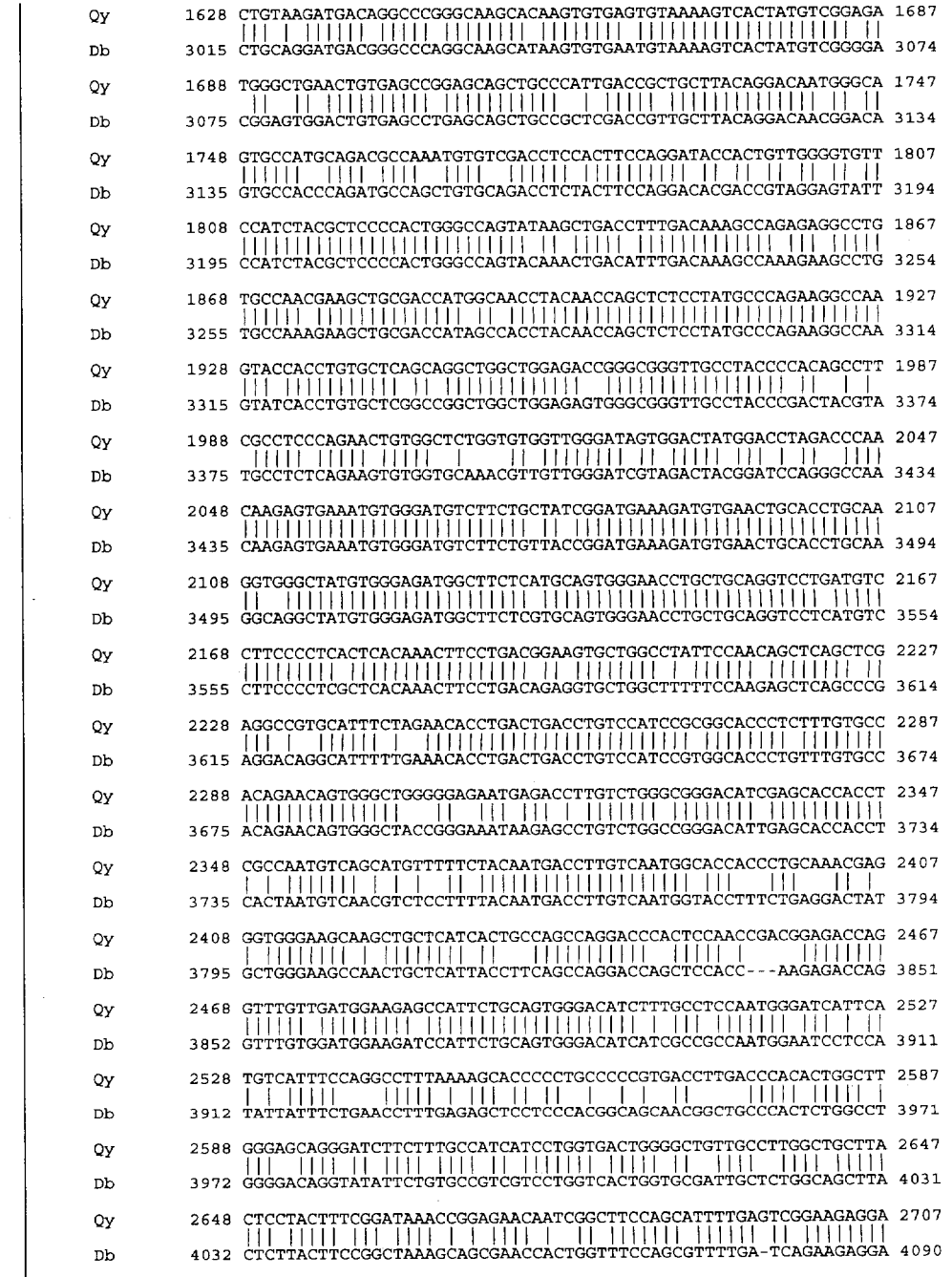
```
LOCUS       AF364951               8157 bp    mRNA    linear   ROD 25-MAR-2002
DEFINITION  Mus musculus stabilin-2 mRNA, complete cds.
ACCESSION   AF364951
VERSION     AF364951.2  GI:19705588
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 8157)
  AUTHORS   Politz,O., Gratchev,A., McCourt,P.A.G., Schledzewski,K.,
            Guillot,P., Johansson,S., Svineng,G., Franke,P., Kannicht,C.,
            Kzhyshkowska,J., Longati,P., Velten,F.W., Johannson,S. and
            Goerdt,S.
  TITLE     Stabilin-1 and -2 constitute a novel family of fasciclin-like
            hyaluronan receptor homologues
  JOURNAL   Biochem. J. 362 (Pt 1), 155-164 (2002)
  MEDLINE   21818465
   PUBMED   11829752
REFERENCE   2  (bases 1 to 8157)
  AUTHORS   Guillot,P., Schledzewski,K., Gratchev,A. and Goerdt,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAR-2001) Dermatology, Universitetsklinikum Mannheim,
            Universitet Heidelberg, Theodor Kutzer Ufer 1-3, Mannheim, BW
            68135, Germany
REFERENCE   3  (bases 1 to 8157)
  AUTHORS   Guillot,P., Schledzewski,K., Gratchev,A. and Goerdt,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-MAR-2002) Dermatology, Universitetsklinikum Mannheim,
            Universitet Heidelberg, Theodor Kutzer Ufer 1-3, Mannheim, BW
            68135, Germany
  REMARK    Sequence update by submitter
COMMENT     On Mar 25, 2002 this sequence version replaced gi:19568177.
FEATURES             Location/Qualifiers
     source          1. .8157
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /tissue_type="liver"
     CDS             62. .7741
                     /note="endothelial protein"
                     /codon_start=1
                     /product="stabilin-2"
                     /protein_id="AAL91684.2"
                     /db_xref="GI:19705589"
                     /translation="MARSKLLLGKLLPLILIFLGLLVQNACSPTEAPELTKRCDKKST
                     LTIKTECQSCSVNIAVKCPDGYIKITNGTVGVRDCRYSLKIQSYVLDIPGCRHICRKD
                     YLQPQCCPGHWGPDCMECPGGARAPCGGRGVCDEGMEGTGSCSCRAGFRGTACENCAA
                     EDVFGPNCSAVCSCVHGVCNSGISGDGTCECLSAYRGPRCDKPIPECAALLCPENSRC
                     SPSSKDETKLQCKCLPSYKGDGQTCKPINPCLKNVCHPHASCSYLGPNRHSCVCQKGY
                     QGDGQVCLPVDPCQTSYGNCPTKSTVCRYDGPGQSHCECKEHYRNFVPGVGCSMTDIC
                     ESKNPCHKNANCSTVSPGQTQCTCQKGYVGDGLNCYGNIMQRLRELNTEPRGMWQGQL
                     TSFISILDRTYAWPLSNLGPFTVLLPSDKGLKGVDVKELLMDKEAARYFVKLHIIAGQ
                     MSTEQMYNLDTFYTLTGKSGEIINKDKDNQLKLKLYGSKIVQIIQGNIVASNGLVHIL
                     DRAMDKIEPTLESNPQQTIMTMLQPRYGKFRSLLEKTNVGQALEKGGIDEPYTIFVPS
                     NEALSNMTAGVLDYLLSPEGSRKLLELVRYHIVAFTQLEVATLVSTLHIRSMANQIIT
                     FNISSKGQILANNVAVDETEVAAKNGRIYTLTGVLIPPSILPILPHRCNETKREMKLG
                     TCVRCFMKNWSKCPTNSEPTAIFTNKCFYGSRAWNLKIGCARYCDVTVEIPRCCKGFF
                     GPDCNPCPGGFMNPCSGNGQCIDGLGGNGTCICEDGFQGSRCQFCSKPNRYGPQCNRT
                     CQCVHGICDNRLDSDGSCLPGTCREGTAGRFCDKQTSACGPYMQFCHIHATCEYSNET
                     ASCVCNDGYEGDGTLCSKKDPCLGSTSRGGCSPNAECIQASTGTYSCVCQRGWTGNGR
                     DCVEINSCLLPSSGGCHDNATCLYVGPGQNECECKKGFRGNGIDCEPIISCLEQIEKC
                     HPLATCQYTLSGVWSCVCQEGYEGNGVLCYGNVLMELSFLSEAAVFYQWINNASLQSM
                     LSATSNLTVLVPSLQAIKDMDQNEKSFWLSRNNIPALIKYHTLLGTYRVADLQTLPSS
                     HMLATSLQGSFLRLDKADGNITIEGASFVDGDNAATNGVVHIINKVLIPQRGLTGSLP
                     SLLTRLEQMPDYSIFRGYIIHYNLASAIEAADAYTVFVPNNEAIESYIREKKATSLKE
                     DILQYHVVLGEKLLRNDLHNGMHRETMLGFSYLLAFFLHNDQLYVNEAPINYTNVATD
                     KGVIHGLEKVLEIKKNRCDNNDTIIVRGKCGKCSQQTLCPLETKPLSETRKCIYSVYF
                     MGKRSIFIGCQLQCVRTIITSACCAGFFGPQCQACPGKGQNVCSGNGFCLDGVNGTGT
                     CECEQGFNGTACETCTEGKYGIHCDQACSCVHGRCNQGPSGDGSCDCDVGWRGVKCDS
                     EITTDNCNGTCHTSANCLLDPDGKASCKCAAGFQGNGTVCTAINACEISNGGCSAKAD
                     CKRTIPGSRVCVCKAGYTGDGIVCLEINPCLENHGGCDRHAECTQTGPNQAVCNCLPK
                     YTGDGKVCTLINVCLTNNGGCSPFAFCNHTEQDQRTCTCKPDYTGDGIVCRGSIHSEL
                     PKNPSTSQYFFQLQEHAVQELAGPGPFTVFVPSSDSFNSESKLKVWDKQGLMSQILRY
                     HVVACQQLLLENLKVITSATTLQGEPISISVSQDTVLINKKAKVLSSDIISTNGVIHV
                     IDTLLSPQNLLITPKGASGRVLLNLTTVAANHGYTKFSKLIQDSGLLKVITDPMHTPV
                     TLFWPTDKALQALPQEQQDFLFNEDNKDKLKAYLKFHVIRDTMALASDLPRSASWKTL
                     QGSELSVRCGTGSDVGELFLNGQMCRIIQRRLLFDGGVAYGIDCLLMDPTEGGRCDTF
                     TTFNIPGECGSCFSTPRCPLQSKPKGVRKKCIYNPLPFRRDVEGCQNLCTLVVHVPRC
                     CSGYFMPDCQACPGGPDTPCNNRGMCYDQYKPTGQCQCHTGFNGTACELCLPGRFGPD
                     CQPCGCSEHGQCDEGITGSGQCLCEAGWTGRFCDAPTVVIPVCIPACSMHATCMENNT
                     CVCNLNYEGDGITCTVVDFCKQNNGGCAKVAKCSQKGTQVSCSCQKGYKGDGHSCTEI
                     DPCANGVNGGCHEHATCRMTGPGKQKCECKSHYVGDGRDCEPEQLPLDRCLQDNGQCH
                     PDANCVDLHFQDTTVGVFHLRSPLGQYKLTFDKAKEACAKEAASIATYNQLSYAQKAK
                     YHLCSAGWLESGRVAYPTIYASKKCANIVGIVDYGTRTNKSEMWDVFCYRMKDVNCTC
                     KAGYVGDGFSCNGNLLQVLMSFPSLTNFLTEVLVFSRSSAQGRAFLKHLTDLSISGTL
                     FVPQNSGLPKNKSLSGRDIEHHLTNVNVSFYDDLVNGTVLKTRLGSQLLITSSQDQLH
                     QEARFVDGRAILQWDIIASNGVLHIISEPLKAPPTAATAAHSGLGTGIFCAVVLVTGA
                     IALAAYSYFRLNQRTTGFRRFESEDDIDALAFGKQQPESITNPLYETSTPAAPEPSCD
                     PFTDSGERELENSDPLGALRS"
ORIGIN
```

```
  Query Match           59.8%;  Score 1949.2;  DB 10;  Length 8157;
  Best Local Similarity  78.0%;  Pred. No. 0;
  Matches 2459;  Conservative    0;  Mismatches  663;  Indels   32;  Gaps    8;

Qy         11 GGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATTTCGT 70
              ||| ||||| ||||| ||   ||| ||||||||||||||||||||||||||||   ||
Db       4885 GGAACTTCCAAAGAATCCCTCAACGTCCCAGTATTTCTTCCAGTTGCAGGAGCATGCTGT 4944

Qy         71 GAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCTTTGA 130
              |||| || |  || || ||||||||||| || |||| |||| ||||| ||||  |
Db       4945 CCAAGAGCTTGCTGGACCTGGCCCCTTCACCGTGTTTGTGCCTTCATCTGACTCCTTCAA 5004

Qy        131 TGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTCGGTA 190
              ||  |   | |||||| |||||| ||| | ||  | ||| ||||| |||| |||||
Db       5005 CAGTGAGTCCAAGCTTAAAGTCTGGGATAAACAGGGCCTCATGTCCCAGATTCTACGGTA 5064

Qy        191 CCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATGCTAC 250
              || ||||| |||||||| ||||| || ||||| ||||| ||  |||||  |  ||| ||
Db       5065 TCACGTGGTGGCCTGCCAGCAGCTACTGCTGGAGAACCTAAAGGTGATCACGAGTGCCAC 5124

Qy        251 TTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAAATAA 310
              ||||||||||||||||||    |||||| ||||||||  ||| |||    ||||| ||
Db       5125 AACCCTCCAAGGAGAGCCAATTTCCATCTCTGTCTCTCAGGACACTGTGCTCATAAACAA 5184

Qy        311 TAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCATAGA 370
              ||||| ||| |  | |||||||| || ||||| || || ||  |  | ||  |||||||
Db       5185 GAAGGCGAAGGTTCTGTCCAGTGACATTATCAGCACCAACGGAGTCATCCACGTCATAGA 5244

Qy        371 CAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAATTCT 430
              || |||| | |||||| |||| |||||||||| |||||||     ||| || ||  ||||
Db       5245 CACATTGTTGTCTCCCCAAAACTTGCTTATCACCCCCAAAGGTGCCTCCGGCAGGGTTCT 5304
```


Qy 431 GCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAATACA 490
Db 5305 GCTAAACCTTACTACAGTGGCAGCAAACCATGGATATACCAAATTCAGCAAGTTGATACA 5364
Qy 491 GGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCTTCTG 550
Db 5365 GGACTCAGGCTTGCTGAAAGTCATCACTGACCCCATGCATACCCCAGTCACTCTCTTCTG 5424
Qy 551 GCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCAACCA 610
Db 5425 GCCTACGGACAAAGCCCTCCAAGCCTTGCCTCAGGAGCAGCAGGACTTCCTGTTCAATGA 5484
Qy 611 AGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCAAGGT 670
Db 5485 AGACAACAAAGACAAGCTGAAGGCATACCTGAAGTTCCACGTGATCCGAGATACCATGGC 5544
Qy 671 TTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGCTGAG 730
Db 5545 TTTAGCTTCAGACCTCCCCAGGTCTGCTTCCTGGAAGACCCTGCAGGGTTCCGAGCTGAG 5604
Qy 731 TGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCTGCAG 790
Db 5605 TGTGAGGTGTGGAACTGGCAGTGATGTCGGAGAGCTCTTTCTAAATGGACAAATGTGCAG 5664
Qy 791 AATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTCTGCT 850
Db 5665 GATTATACAGCGGAGACTCTTGTTTGACGGAGGTGTGGCCTATGGCATTGACTGCCTGCT 5724
Qy 851 GATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGGGGGA 910
Db 5725 AATGGATCCTACCGAAGGTGGCCGATGTGACACTTTTACTACCTTCAATATTCCGGGGGA 5784
Qy 911 GTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTGTGAA 970
Db 5785 ATGTGGGAGCTGTTTCTCCACTCCCAGATGCCCACTACAGAGCAAACCAAAGGGTGTGAG 5844
Qy 971 GCAGAAGTGTCTCTACAACC---TGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGCG 1027
Db 5845 GAAGAAGTGTATCTACAACCCGTTACCTTTCAGGAGGGATGTGGAAGGCTGTCAGAACCT 5904
Qy 1028 GTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTCA 1087
Db 5905 GTGCACCCTGGTGGTCCATGTCCCGCGGTGCTGCAGTGGTTACTTCATGCCAGACTGTCA 5964
Qy 1088 GGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGTA 1147
Db 5965 GGCCTGCCCTGGAGGACCAGACACACCGTGTAACAACCGGGGCATGTGTTATGATCAGTA 6024
Qy 1148 CTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGTG 1207
Db 6025 CAAACCCACAGGACAGTGCCAATGTCATACTGGTTTCAACGGGACAGCCTGTGAGCTCTG 6084
Qy 1208 CTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGTG 1267
Db 6085 CTTGCCTGGGAGATTTGGTCCTGACTGTCAACCCTGTGGCTGCTCTGAGCATGGGCAGTG 6144
Qy 1268 CGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCTC 1327
Db 6145 TGACGAGGGGATCACAGGCTCCGGGCAGTGCCTCTGTGAAGCAGGGTGGACAGGCCGCTT 6204
Qy 1328 GTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCAC 1387
Db 6205 CTGTGATGCTCCCACAGTTGTGATTCCAGTGTGCATACCTGCTTGTTCCATGCACGCAAC 6264
Qy 1388 CTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACATG 1447
Db 6265 CTGTATGGAGAACAACACATGTGTGTGTAACTTGAATTACGAAGGTGACGGAATCACATG 6324
Qy 1448 CACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCTC 1507
Db 6325 CACAGTCGTGGATTTCTGCAAACAGAACAATGGGGGCTGTGCGAAGGTCGCCAAGTGCTC 6384
Qy 1508 CCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACAG 1567
Db 6385 CCAGAAGGGCACACAAGTCTCCTGCAGCTGCCAGAAGGGCTACAAGGGGGATGGCCACAG 6444
Qy 1568 CTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCAC 1627
Db 6445 CTGCACAGAGATAGACCCCTGTGCAAACGGTGTCAATGGTGGATGTCATGAGCATGCTAC 6504
Qy 1628 CTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAGA 1687
Db 6505 CTGCAGGATGACGGGCCCAGGCAAGCAGAAGTGTGAATGTAAAAGTCACTATGTCGGGGA 6564
Qy 1688 TGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGCA 1747
Db 6565 TGGACGGGACTGTGAGCCCGAACAGCTGCCCCTTGACCGCTGCTTACAGGACAACGGACA 6624
Qy 1748 GTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGTT 1807
Db 6625 ATGCCACCCAGATGCCAACTGTGTGGACCTTCATTTCCAGGACACTACTGTTGGCGTATT 6684
Qy 1808 CCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCTG 1867
Db 6685 CCACCTACGCTCCCCACTGGGCCAGTACAAACTGACATTTGACAAAGCCAAAGAGGCCTG 6744
Qy 1868 TGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAA 1927
Db 6745 TGCTAAGGAAGCTGCATCCATAGCCACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAA 6804
Qy 1928 GTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCTT 1987
Db 6805 GTATCACCTCTGCTCAGCTGGCTGGCTGGAGAGTGGGCGGGTTGCCTACCCCACGATCTA 6864
Qy 1988 CGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCAA 2047
Db 6865 CGCCTCTAAGAAATGTG---CAAACATTGTGGGGATTGTAGACTACGGAACCAGGACCAA 6921
Qy 2048 CAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCAA 2107
Db 6922 CAAGAGTGAAATGTGGGACGTCTTCTGTTACAGAATGAAAGATGTTAACTGCACCTGCAA 6981
Qy 2108 GGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGTC 2167
Db 6982 GGCAGGCTACGTGGGCGATGGCTTCTCCTGCAATGGGAACCTGCTGCAAGTCCTCATGTC 7041
Qy 2168 CTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTCG 2227
Db 7042 CTTCCCCTCACTCACGAACTTCCTGACAGAGGTGCTGGTTTTTTCCAGGAGCTCAGCCCA 7101
Qy 2228 AGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCC 2287
Db 7102 AGGCCGGGCGTTTTTGAAACACCTGACTGACCTGTCCATCAGTGGCACCCTGTTTGTGCC 7161
Qy 2288 ACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACCT 2347
Db 7162 ACAGAACAGTGGGCTACCGAAAAATAAGAGCCTGTCTGGGCGGGACATTGAGCACCACCT 7221
Qy 2348 CGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGAG 2407
Db 7222 CACTAATGTCAACGTCTCCTTTTACGATGACCTTGTCAATGGTACCGTCCTGAAGACTAG 7281
Qy 2408 GGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCAG 2467
Db 7282 GCTGGGAAGCCAACTGCTCATCACCTCCAGCCAGGACCAGCTCCACC---AAGAGGCCAG 7338
Qy 2468 GTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCA 2527
Db 7339 GTTCGTGGATGGAAGAGCCATTCTGCAGTGGGACATCATTGCCTCTAACGGGGTTCTCCA 7398
Qy 2528 TGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCTT 2587
Db 7399 TATCATTTCTGAACCTTTGAAAGCTCCTCCCACGGCCGCAACGGCTGCCCACTCTGGCCT 7458
Qy 2588 GGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTTA 2647

```
Db       7459 AGGAACAGGCATATTCTGCGCTGTCGTCCTGGTTACTGGTGCGATTGCTCTGGCTGCCTA 7518

Qy       2648 CTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGA 2707

Db       7519 CTCCTACTTCCGGCTAAACCAGAGAACAACTGGCTTCCGGCGTTTTGAGTCAGAAGATGA 7578

Qy       2708 CATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGA 2767

Db       7579 CATTGATGCCTTGGCCTTTGGCAAGCAGCAGCCTGAGAGTATCACAAACCCTTTGTATGA 7638

Qy       2768 GAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACG 2827

Db       7639 GACCTCAACGCCGGCAGCCCCCGAGCCCTCCTGTGACCCCTTCACCGACTCTGGAGAACG 7698

Qy       2828 GCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGC 2887

Db       7699 GGAGCTGGAGAACAGTGACCCTCTTGGGGCACTG---AGGTCCTGATGTGAGAAGCCAGC 7755

Qy       2888 CATCACTCACTGCCACCTGGGCCATCA----------ACTGTGAATTCTCAGCACC 2933

Db       7756 CAACAGCCACAGCCACGTGAGCCCTGAGCCATCACGGTTACACGGTGACTCCCCTGCTCC 7815

Qy       2934 AGTTGCCTTTTAG---GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCT 2990

Db       7816 AGCTGTCACTTGGATCATTTGTTTTAAAGTATGATAACTCAGAAGCCATACCTCACCCCT 7875

Qy       2991 CTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTAC 3050

Db       7876 CTGGTTAGTCTGGGGTTGTCGTCATGGGTAAGGGGACATGTTTCCAGGACACCAGTATAC 7935

Qy       3051 AGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTC 3110

Db       7936 CTCTGCCTCCTCTGGGCCTTCACTGTGGCT--CTCTGCACCTTCCAGATGGTGCTTGGCC 7993

Qy       3111 CATTCTGCCCTACATGATGGGTAACTGTGATCTT 3144

Db       7994 TGTTCTGCCCTATCTG-TGGGTAGCTGTGACCCT 8026


RESULT 14
AY007370
LOCUS       AY007370                4706 bp    mRNA    linear   ROD 23-OCT-2002
DEFINITION  Rattus norvegicus hyaluronan receptor for endocytosis HARE
            precursor, mRNA, partial cds.
ACCESSION   AY007370
VERSION     AY007370.1  GI:24285892
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 4706)
  AUTHORS   Zhou,B., Weigel,J.A., Saxena,A. and Weigel,P.H.
  TITLE     Molecular Cloning and Functional Expression of the Rat 175-kDa
            Hyaluronan Receptor for Endocytosis
  JOURNAL   Mol. Biol. Cell 13 (8), 2853-2868 (2002)
  MEDLINE   22169209
   PUBMED   12181351
REFERENCE   2  (bases 1 to 4706)
  AUTHORS   Zhou,B., Weigel,J.A. and Weigel,P.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-AUG-2000) Department of Biochemistry and Molecular
            Biology, University of Oklahoma Health Sciences Center, 940 Stanton
            L. Young, BMSB 860, Oklahoma City, OK 73104, USA
FEATURES             Location/Qualifiers
     source          1. .4706
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
     CDS             <1. .4296
                     /codon_start=1
                     /product="hyaluronan receptor for endocytosis HARE
                     precursor"
                     /protein_id="AAG13634.1"
                     /db_xref="GI:24285893"
                     /translation="SLPSLLTRLEQMPDYSIFRGYIIHYNLASAIESADAYTVFVPNN
                     EAIENYIREKKATSLKEDILRYHVVLGEKLLKNDLHNGMHRETMLGFSYLLAFFLRND
                     QLYVNEAPINYTNVATDKGVIHGLEKVLEIQKNRCDNNDTIIVRGECGKCSQQAPCPL
                     ETKPLRETRKCIYSIYFMGKRSVFIGCQPQCVRTIITRACCAGFFGPQCQACPGRGQN
                     VCSGNGFCLDGVNGTGTCQCGLGFNGTACETCTEGKYGIHCDQACSCVHGRCSQGPLG
                     DGSCDCDVGWRGVKCDMEITTDNCNGTCHTSANCLLDPDGKASCKCAAGFRGNGTVCT
                     AINACETSNGGCSTKADCKRTTPGNRVCVCKAGYTGDGIVCLEINPCLENHGGCDRNA
                     ECTQTGPNQAVCNCLPKYTGDGKVCSLINVCLTNNGGCSPFAFCNYTEQDQRICTCKP
                     DYTGDGIVCRGSIYGELPKNPSTSQYFFQLQEHAVRELAGPGPFTVFAPLSSSFNHEP
                     RIKDWDQQGLMSQVLRYHVVGCQQLLLDNLKVTTSATTLQGEPVSISVSQDTVFINNE
                     AKVLSSDIISTNGVIHVIDKLLSPKNLLITPKDALGRVLQNLTTVAANHGYTKFSKLI
                     QDSGLLSVITDSIHTPVTVFWPTDKALEALPPEQQDFLFNQDNKDKLKSYLKFHVIRD
                     SKALASDLPRSASWKTLQGSELSVRCGTGSDIGELFLNEQMCRFIHRGLLFDVGVAYG
                     IDCLLMNPTLGGRCDTFTTFDIPGECGSCIFTPKCPLKSKPKGVKKKCIYNPLPFRRN
                     VEGCQNLCTVVIQTPRCCHGYFMPDCQACPGGPDTPCNNRGMCRDLYTPMGQCLCHTG
                     FNGTACELCWHGRFGPDCQPRSCSEHGQCDEGITGSGECLCETGWTAASCDTPTAVFA
                     VCTPACSVHATCTENNTCVCNLNYEGDGITCTVVDFCKQNNGGCAKVAKCSQKGTQVS
                     CSCKKGYKGDGYSCIEIDPCADGVNGGCHEHATCRMTGPGKHKCECKSHYVGDGVDCE
                     PEQLPLDRCLQDNGQCHPDASCADLYFQDTTVGVFHLRSPLGQYKLTFDKAKEACAKE
                     AATIATYNQLSYAQKAKYHLCSAGWLESGRVAYPTTYASQKCGANVVGIVDYGSRANK
                     SEMWDVFCYRMKDVNCTCKAGYVGDGFSCSGNLLQVLMSFPSLTNFLTEVLAFSKSSA
                     RGQAFLKHLTDLSIRGTLFVPQNSGLPGNKSLSGRDIEHHLTNVNVSFYNDLVNGTFL
                     RTMLGSQLLITFSQDQLHQETRFVDGRSILQWDIIAANGILHIISEPLRAPPTAATAA
                     HSGLGTGIFCAVVLVTGAIALAAYSYFRLKQRTTGFQRFDQKRTLMSWLLASSSPRIS
                     QTLCMRPQRRHPQSPPVTPSQTLENRIWRTATLWGHCGPDMRSQQATTVTVPR"
     mat_peptide     1. .4293
                     /product="hyaluronan receptor for endocytosis HARE"
                     /note="175 kDa; derived by proteolytic cleavage"
ORIGIN

  Query Match           59.4%;  Score 1936.8;  DB 10;  Length 4706;
  Best Local Similarity   78.3%;  Pred. No. 0;
  Matches 2402;  Conservative    0;  Mismatches  652;  Indels   14;  Gaps    6;

Qy         11 GGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATTTCGT 70

Db       1395 GGAGCTTCCCAAGAACCCTTCGACGTCCCAGTACTTCTTCCAGTTGCAGGAGCATGCTGT 1454

Qy         71 GAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCTTTGA 130

Db       1455 CCGAGAGCTTGCTGGACCTGGCCCCTTCACCGTGTTCGCGCCTTTGTCTAGCTCCTTCAA 1514

Qy        131 TGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTCGGTA 190

Db       1515 TCATGAGCCCCGGATTAAAGACTGGGATCAGCAGGGCCTCATGTCCCAGGTTCTTCGCTA 1574

Qy        191 CCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATGCTAC 250

Db       1575 TCACGTGGTGGGCTGCCAGCAGCTGCTGTTGGACAACCTAAAAGTGACCACAAGTGCCAC 1634

Qy        251 TTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAAATAA 310

Db       1635 GACCCTCCAAGGAGAGCCAGTTTCCATCTCTGTCTCTCAGGACACTGTGTTCATAAACAA 1694

Qy        311 TAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCATAGA 370

Db       1695 TGAGGCGAAGGTCCTGTCCAGTGACATCATCAGCACCAATGGCGTCATCCACGTTATAGA 1754

Qy        371 CAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAATTCT 430

Db       1755 CAAGTTGCTGTCTCCCAAAAACTTGCTTATCACCCCCAAAGATGCCTTGGGCAGGGTTCT 1814

Qy        431 GCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAATACA 490

Db       1815 GCAAAATCTTACTACAGTGGCAGCAAACCACGGATATACCAAATTCAGCAAGTTGATACA 1874

Qy        491 GGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCTTCTG 550

Db       1875 GGACTCAGGCTTGCTGTCAGTCATCACTGACTCCATCCACACCCCAGTCACTGTCTTCTG 1934
```


Qy 551 GCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCAACCA 610
Db 1935 GCCTACGGACAAAGCCCTGGAAGCCTTGCCCCCAGAGCAGCAGGACTTCCTGTTCAATCA 1994
Qy 611 AGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCAAGGT 670
Db 1995 AGACAACAAGGACAAGCTGAAGTCTTACCTGAAGTTCCACGTGATCCGAGACTCCAAGGC 2054
Qy 671 TTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGCTGAG 730
Db 2055 TTTAGCTTCAGACCTCCCCAGGTCTGCTTCCTGGAAGACCCTGCAAGGCTCAGAGCTGAG 2114
Qy 731 TGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCTGCAG 790
Db 2115 TGTGAGGTGTGGAACTGGCAGTGACATCGGTGAGCTCTTTCTAAACGAACAAATGTGCAG 2174
Qy 791 AATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTCTGCT 850
Db 2175 ATTCATACACCGGGGACTCTTGTTTGACGTGGGTGTGGCCTATGGCATTGACTGCCTACT 2234
Qy 851 GATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGGGGGA 910
Db 2235 CATGAATCCTACCCTAGGTGGCCGATGTGACACTTTTACTACCTTCGATATTCCGGGGGA 2294
Qy 911 GTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTGTGAA 970
Db 2295 GTGCGGAAGTTGCATTTTCACTCCCAAATGCCCACTGAAGAGCAAGCCAAAGGGCGTGAA 2354
Qy 971 GCAGAAGTGTCTCTACAACC---TGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGCG 1027
Db 2355 GAAGAAGTGTATCTACAACCCGTTACCTTTCAGGAGGAACGTGGAAGGCTGCCAGAACCT 2414
Qy 1028 GTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTCA 1087
Db 2415 GTGCACCGTGGTGATCCAAACCCCCAGGTGCTGCCATGGTTACTTCATGCCAGACTGTCA 2474
Qy 1088 GGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGTA 1147
Db 2475 GGCCTGCCCTGGAGGACCAGATACACCGTGTAACAACCGGGGCATGTGCCGCGATCTGTA 2534
Qy 1148 CTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGTG 1207
Db 2535 CACACCCATGGGACAGTGCCTATGCCACACCGGCTTCAACGGGACAGCCTGCGAGCTCTG 2594
Qy 1208 CTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGTG 1267
Db 2595 CTGGCATGGGAGATTTGGGCCTGACTGTCAGCCCCGCAGCTGCTCCGAGCATGGACAGTG 2654
Qy 1268 CGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCTC 1327
Db 2655 TGATGAGGGGATCACAGGCTCCGGGGAGTGCCTCTGTGAAACAGGGTGGACAGCCGCTTC 2714
Qy 1328 GTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCAC 1387
Db 2715 GTGTGACACTCCCACAGCTGTATTCGCAGTGTGCACACCTGCTTGCTCCGTGCACGCCAC 2774
Qy 1388 CTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACATG 1447
Db 2775 CTGTACGGAGAACAACACGTGTGTGTGTAACTTGAACTACGAAGGTGACGGGATCACATG 2834
Qy 1448 CACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCTC 1507
Db 2835 CACAGTCGTGGACTTCTGCAAACAGAACAACGGGGGCTGTGCGAAGGTCGCTAAGTGCTC 2894
Qy 1508 CCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACAG 1567
Db 2895 CCAGAAAGGCACCCAAGTCTCTTGCAGCTGCAAGAAAGGCTACAAGGGGGATGGCTACAG 2954
Qy 1568 CTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCAC 1627
Db 2955 CTGCATAGAGATAGACCCCTGTGCAGACGGTGTCAACGGGGGATGCCATGAGCACGCCAC 3014


Qy 1628 CTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAGA 1687
Db 3015 CTGCAGGATGACGGGCCCAGGCAAGCATAAGTGTGAATGTAAAAGTCACTATGTCGGGGA 3074
Qy 1688 TGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGCA 1747
Db 3075 CGGAGTGGACTGTGAGCCTGAGCAGCTGCCGCTCGACCGTTGCTTACAGGACAACGGACA 3134
Qy 1748 GTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGTT 1807
Db 3135 GTGCCACCCAGATGCCAGCTGTGCAGACCTCTACTTCCAGGACACGACCGTAGGAGTATT 3194
Qy 1808 CCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCTG 1867
Db 3195 CCATCTACGCTCCCCACTGGGCCAGTACAAACTGACATTTGACAAAGCCAAAGAAGCCTG 3254
Qy 1868 TGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAA 1927
Db 3255 TGCCAAAGAAGCTGCGACCATAGCCACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAA 3314
Qy 1928 GTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCTT 1987
Db 3315 GTATCACCTGTGCTCGGCCGGCTGGCTGGAGAGTGGGCGGGTTGCCTACCCGACTACGTA 3374
Qy 1988 CGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCAA 2047
Db 3375 TGCCTCTCAGAAGTGTGGTGCAAACGTTGTTGGGATCGTAGACTACGGATCCAGGGCCAA 3434
Qy 2048 CAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCAA 2107
Db 3435 CAAGAGTGAAATGTGGGATGTCTTCTGTTACCGGATGAAAGATGTGAACTGCACCTGCAA 3494
Qy 2108 GGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGTC 2167
Db 3495 GGCAGGCTATGTGGGAGATGGCTTCTCGTGCAGTGGGAACCTGCTGCAGGTCCTCATGTC 3554
Qy 2168 CTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTCG 2227
Db 3555 CTTCCCCTCGCTCACAAACTTCCTGACAGAGGTGCTGGCTTTTTCCAAGAGCTCAGCCCG 3614
Qy 2228 AGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCC 2287
Db 3615 AGGACAGGCATTTTTGAAACACCTGACTGACCTGTCCATCCGTGGCACCCTGTTTGTGCC 3674
Qy 2288 ACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACCT 2347
Db 3675 ACAGAACAGTGGGCTACCGGGAAATAAGAGCCTGTCTGGCCGGGACATTGAGCACCACCT 3734
Qy 2348 CGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGAG 2407
Db 3735 CACTAATGTCAACGTCTCCTTTTACAATGACCTTGTCAATGGTACCTTTCTGAGGACTAT 3794
Qy 2408 GGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCAG 2467
Db 3795 GCTGGGAAGCCAACTGCTCATTACCTTCAGCCAGGACCAGCTCCACC---AAGAGACCAG 3851
Qy 2468 GTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCA 2527
Db 3852 GTTTGTGGATGGAAGATCCATTCTGCAGTGGGACATCATCGCCGCCAATGGAATCCTCCA 3911
Qy 2528 TGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCTT 2587
Db 3912 TATTATTTCTGAACCTTTGAGAGCTCCTCCCACGGCAGCAACGGCTGCCCACTCTGGCCT 3971
Qy 2588 GGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTTA 2647
Db 3972 GGGGACAGGTATATTCTGTGCCGTCGTCCTGGTCACTGGTGCGATTGCTCTGGCAGCTTA 4031
Qy 2648 CTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGA 2707
Db 4032 CTCTTACTTCCGGCTAAAGCAGCGAACCACTGGTTTCCAGCGTTTTGA-TCAGAAGAGGA 4090
Qy 2708 CATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGA 2767

```
Db       4091 CATTGATGTCTTGGCTTTTGGCAAGCAGCAGCCCAAGAATATCGCAAACCCTCTGTATGA 4150

Qy       2768 GAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACG 2827

Db       4151 GACCTCAGCGCCGGCACCCCCAGAGTCCTCCTGTGACCCCTTCACAGACCCTGGAGAACA 4210

Qy       2828 GCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGC 2887

Db       4211 GGATCTGGAGGACAGCGACCCTCTGGGGGCACTG---CGGTCCTGACATGAGAAGCCAGC 4267

Qy       2888 CATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGC---ACCAGTTGCCTTTT 2944

Db       4268 AAGCAACCACAGTCACGGTTCCACGGTGATTCCCAGCCCCAGCTGTCTCATGGATCAGTT 4327

Qy       2945 AGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGG 3004

Db       4328 GTTTTAAAGAATGACAACACTCATAAGCCAGCCATACCTCACCCTTCTGGTTAATCTGGG 4387

Qy       3005 GGTTGTTTCTGTGGGTGAGAGATGTGTTG-CTGTGCCCACCCAGTACAGCTTCCTCCTCT 3063

Db       4388 ATTGTCGCCAGGGCTAAGGAGCCATGTTGCCTGGATACCTGGGGGACCTCCACCTCCTCT 4447

Qy       3064 GACCCTTT 3071

Db       4448 GAGCCTAT 4455


RESULT 15
AX149459
LOCUS       AX149459               2483 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 15 from Patent WO0136638.
ACCESSION   AX149459
VERSION     AX149459.1  GI:14347935
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A., Lichenstein,H., Vernet,C. and Fernandes,E.
  TITLE     Polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0136638-A 15 25-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .2483
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           54.6%;  Score 1778.6;  DB 6;  Length 2483;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1790;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy        122 AGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGT 181

Db        569 ACCCTCTGATCTTCTGTTACAGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGT 628

Qy        182 TCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTC 241

Db        629 TCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTC 688

Qy        242 AAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTA 301

Db        689 AAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTA 748

Qy        302 TATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCA 361

Db        749 TATAAACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCA 808

Qy        362 TATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGG 421

Db        809 TATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGG 868

Qy        422 AAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAA 481

Db        869 AAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAA 928

Qy        482 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCAC 541

Db        929 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCAC 988

Qy        542 TCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCT 601

Db        989 TCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCT 1048

Qy        602 GTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 661

Db       1049 GTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 1108

Qy        662 TGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTC 721

Db       1109 TGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTC 1168

Qy        722 AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCA 781

Db       1169 AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCA 1228

Qy        782 AACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGA 841

Db       1229 AACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGA 1288

Qy        842 CTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGC 901

Db       1289 CTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGC 1348

Qy        902 CTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAA 961

Db       1349 CTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAA 1408

Qy        962 GGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCG 1021

Db       1409 GGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCG 1468

Qy       1022 GGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGA 1081

Db       1469 GGAGCGATGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGA 1528

Qy       1082 CTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGA 1141

Db       1529 CTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGA 1588

Qy       1142 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA 1201

Db       1589 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA 1648

Qy       1202 GATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGG 1261

Db       1649 GATGTGCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGG 1708

Qy       1262 ACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGG 1321

Db       1709 ACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGG 1768

Qy       1322 CCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCA 1381

Db       1769 CCCCTCGTGTGACACTCAGGCAGTTTTGTCTGCAGTGTGTACGCCTCCTTGTTCTGCTCA 1828

Qy       1382 TGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAAT 1441

Db       1829 TGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAAT 1888

Qy       1442 CACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAG 1501

Db       1889 CACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAG 1948
```

```
Qy       1502 ATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGG 1561
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1949 ATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGG 2008

Qy       1562 GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCA 1621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2009 GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCA 2068

Qy       1622 CGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGT 1681
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2069 CGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGT 2128

Qy       1682 CGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAA 1741
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2129 CGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAA 2188

Qy       1742 TGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGG 1801
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2189 TGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGG 2248

Qy       1802 GGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA 1861
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2249 GGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA 2308

Qy       1862 GGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAA 1921
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2309 GGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAA 2368

Qy       1922 GGCCAAGTA 1930
              | ||  |||
Db       2369 GACCTGGTA 2377
```

Search completed: June 21, 2004, 14:12:32
Job time : 12320 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 10:23:51 ; Search time 216 Seconds
(without alignments)
8375.648 Million cell updates/sec

Title: US-10-028-248A-210
Perfect score: 3260
Sequence: 1 ggcacgagcaggagcttccc.....aaaaaaaaaaaaaaaaaaa 3260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	187.4	5.7	1482	4	US-09-016-434-230	Sequence 230, App
	2	107.2	3.3	409	4	US-09-833-381-1252	Sequence 1252, Ap
	3	81.4	2.5	1414	1	US-08-024-868-1	Sequence 1, Appli
	4	81.4	2.5	1414	2	US-08-242-097-1	Sequence 1, Appli
	5	81.4	2.5	1414	3	US-09-206-695-1	Sequence 1, Appli
	6	81.4	2.5	1414	4	US-09-799-118-1	Sequence 1, Appli
	7	72	2.2	1734	4	US-09-484-970B-63	Sequence 63, Appl
	8	58.6	1.8	1985	4	US-09-907-794A-212	Sequence 212, App
	9	58.6	1.8	1985	4	US-09-905-125A-212	Sequence 212, App
	10	58.6	1.8	1985	4	US-09-902-775A-212	Sequence 212, App
c	11	54.2	1.7	7218	1	US-08-232-463-14	Sequence 14, Appl
	12	53.8	1.7	1400	2	US-08-001-078A-2	Sequence 2, Appli
	13	53.8	1.7	1400	2	US-08-463-218-2	Sequence 2, Appli
	14	53.8	1.7	1400	5	PCT-US94-00253-2	Sequence 2, Appli
	15	52.8	1.6	7218	1	US-08-232-463-14	Sequence 14, Appl
	16	47.6	1.5	3259	5	PCT-US95-03747-1	Sequence 1, Appli
	17	44.4	1.4	1587	4	US-09-010-147B-19	Sequence 19, Appl
	18	44.4	1.4	1720	4	US-09-148-545-53	Sequence 53, Appl
	19	43.2	1.3	5191	1	US-08-340-428B-1	Sequence 1, Appli
	20	43.2	1.3	5191	5	PCT-US93-07306-1	Sequence 1, Appli
	21	42.2	1.3	2148	4	US-09-643-597-154	Sequence 154, App
	22	42.2	1.3	2148	4	US-09-480-884A-154	Sequence 154, App
	23	42.2	1.3	2148	4	US-09-542-615A-154	Sequence 154, App
	24	42.2	1.3	2148	4	US-09-606-421B-154	Sequence 154, App
	25	42.2	1.3	2148	4	US-09-221-107-154	Sequence 154, App
	26	41.8	1.3	619	4	US-09-489-847-58	Sequence 58, Appl
	27	41	1.3	94	4	US-09-621-976-12471	Sequence 12471, A
	28	41	1.3	8224	6	5180808-1	Patent No. 5180808
	29	40.8	1.3	1520	1	US-08-225-477B-1	Sequence 1, Appli
	30	40.8	1.3	1520	5	PCT-US95-04353-1	Sequence 1, Appli
c	31	39.6	1.2	832	4	US-09-621-976-2813	Sequence 2813, Ap
	32	39.4	1.2	855	6	5185441-40	Patent No. 5185441
	33	39.4	1.2	855	6	5223394-3	Patent No. 5223394
	34	39.4	1.2	863	1	US-07-940-861-11	Sequence 11, Appl
	35	39.4	1.2	863	1	US-08-459-512-11	Sequence 11, Appl
	36	39.4	1.2	863	2	US-08-459-657-11	Sequence 11, Appl
	37	39.4	1.2	863	2	US-08-460-132-11	Sequence 11, Appl
	38	39.4	1.2	863	5	PCT-US92-02050-11	Sequence 11, Appl
	39	39.4	1.2	863	6	5185441-35	Patent No. 5185441
	40	39.4	1.2	863	6	5223394-5	Patent No. 5223394
	41	39.4	1.2	1078	6	5223394-10	Patent No. 5223394
	42	39.4	1.2	1519	1	US-08-225-477B-2	Sequence 2, Appli
	43	39.4	1.2	1519	5	PCT-US95-04353-2	Sequence 2, Appli
	44	39.4	1.2	2007	4	US-09-643-597-153	Sequence 153, App
	45	39.4	1.2	2007	4	US-09-480-884A-153	Sequence 153, App

ALIGNMENTS

```
RESULT 1
US-09-016-434-230
; Sequence 230, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  230:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1482 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SPLNNOT04
;      CLONE:  1561663
US-09-016-434-230

  Query Match             5.7%;  Score 187.4;  DB 4;  Length 1482;
  Best Local Similarity   53.2%;  Pred. No. 4.3e-45;
```

```
  Matches  493;  Conservative    0;  Mismatches  421;  Indels   13;  Gaps    4;

Qy       1659 TGTGAGTGTAAAAGTCACTATGTCGGAGATGGGCTGAACTGT---GAGCCGGAGCAGCTG 1715
             ||||||||  |       ||| || || ||||| ||| | |||    |||  |  | | |
Db          7 TGTGAGTGCCACGCAGGCTACGTAGGCGATGGACTGCAGTGTCTGGAGGAGTCGGAACCA 66

Qy       1716 CCCATTGACCGCTGCTTACAGGACAATGGGCAGTGCCATGCAGACGCCAAATGTGTCGAC 1775
             ||  | ||||||||||  |     ||  |||||  |||| ||||  ||     |||
Db         67 CCTGTGGACCGCTGCTTGGGCCAGCCACCGCCCTGCCACTCAGATGCCATGTGCACTGAC 126

Qy       1776 CTCCACTTCCAGGATACCACTGTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTAT 1835
             || |||||||||| |     | ||| || ||||| || |   || |    ||||  |||
Db         127 CTGCACTTCCAGGAGAAACGGGCTGGCGTTTTCCACCTCCAGGCCACCAGCGGCCCTTAT 186

Qy       1836 AAGCTGACCTTTGACAAAGCCAGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACC 1895
                |||| ||||     | ||     | ||| || |   || | |  ||    | | ||  |
Db         187 GGTCTGAACTTTTCGGAGGCTGAGGCGGCATGCGAAGAC-AGGGAGCCGTCCTTGCTTCA 245

Qy       1896 TACAACCAGCTCTCCTATGCCCAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTG 1955
             | |  ||||||||  |||||||| ||      | ||||||||||      |||||||||
Db         246 TTCCCTCAGCTCTCTGCTGCCCAGCAGCTGGGCTTCCACCTGTGCCTCATGGGCTGGCTG 305

Qy       1956 GAGACCGGGCGGGTTGCCTACCCCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTG 2015
             |   |  ||       |||| ||||     |  ||| |     | ||||||| ||||  |
Db         306 GCCAATGGCTCCACTGCCCACCCTGTGGTTTTCCCTGTGGCGGACTGTGGCAATGGTCGG 365

Qy       2016 GTTGGGATAGTGGACTATGGACCTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGC 2075
             || || |||||    |   ||  |  |    |||      |||   ||||||| || ||||
Db         366 GTGGGCATAGTCAGCCTGGGTGCCCGCAAGAACCTCTCAGAACGCTGGGATGCCTACTGC 425

Qy       2076 TATCGGATGAAAGATGTGAACTGCACCTGCAAGGTGGGCTATGTGGGAGATGG---CTTC 2132
             |  ||  || ||||||||  ||||   |||     ||||  ||||| || ||    |  |
Db         426 TTCCGTGTGCAAGATGTGGCCTGCCGATGCCGAAATGGCTTCGTGGGTGACGGGATCAGC 485

Qy       2133 TCATGCAGTGGGAACCTGCTGCAGGTCCTGATGTCCTTCCCCTCACTCACAAACTTCCTG 2192
               |||| ||||||  ||||||  | || |||     ||     ||  || | | ||||
Db         486 ACGTGCAATGGGAAGCTGCTGGATGTGCTGGCTGCCACTGCCAACTTCTCCACCTTCTAT 545

Qy       2193 ACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTG 2252
                |   | |||  |||| ||||   | |    || || |  |  || || ||   ||||
Db         546 GGGATGCTATTGGGCTATGCCAATGCCACCCAGCGGGGTCTCGACTTCCTGGACTTCCTG 605

Qy       2253 ACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAAT 2312
                |||   |  ||          || ||||| || ||    ||     ||  | | ||| ||
Db         606 GATGATGAGCTCACGTATAAGACACTCTTCGTCCCTGTCAATGAAGGCTTTGTGGACAAC 665

Qy       2313 GAGACCTTGTCTGGGCGGGACATCGAGCACCACCTCGCCAATGTCAGCATGTTTTTCTAC 2372
                |||  ||  |||  |  |||  | ||||  ||   | ||||  | ||  |  |     |
Db         666 ATGACGCTGAGTGGCCCAGACTTGGAGCTGCATGCCTCCAACGCCACCCTCCTAAGTGCC 725

Qy       2373 AATGACCTTGTCAATGGCACCACCCTGCAAACGAGGGTGGGAAGCAAGCTGCTCATCACT 2432
             ||  | |                | |         |     |||     |  |||
Db         726 AACGCCAGCCAGGGGAAGTTGCTTCCGGCCCACTCAGGCCTCAGCCTCATCATCAGTGAC 785

Qy       2433 GCCAGCCAGGACC------CACTCCAACCGACGGAGACCAGGTTTGTTGATGGAAGAGCC 2486
             ||  |||  |||           |         ||   ||   |  ||   ||||  |  ||
Db         786 GCAGGCCCTGACAACAGTTCCTGGGCCCCTGTGGCCCCAGGGACAGTTGTGGTTAGCCGT 845

Qy       2487 ATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCATGTCATTTCCAGGCCTTTA 2546
             ||  |  |||||||||| | |||| ||||||  |||||  ||||    |  ||||  ||  |
Db         846 ATCATTGTGTGGGACATCATGGCCTTCAATGGCATCATCCATGCTCTGGCCAGCCCCCTC 905

Qy       2547 AAAGCACCCCCTGCCCCCGTGACCTTG 2573
                 ||||||||    |||  |  |  ||
Db         906 CTGGCACCCCCACAGCCCCAGGCAGTG 932


RESULT 2
US-09-833-381-1252
; Sequence 1252, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1252
;   LENGTH: 409
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(409)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1252

  Query Match             3.3%;  Score 107.2;  DB 4;  Length 409;
  Best Local Similarity   56.3%;  Pred. No. 1.2e-21;
  Matches  218;  Conservative    0;  Mismatches  166;  Indels    3;  Gaps    1;

Qy       1452 GTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCTCCCAG 1511
             || | |||  | ||| |  |||   | || ||||||     |||   ||||  |||  ||||
Db         23 GTGGCGGANCTGTGCCAGAAGGGGCATGGTGGCTGCAGTAAGCATGCCAACTGCAGCCAG 82

Qy       1512 AAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACAGCTGC 1571
                ||| ||   |||    |||| |||||  |   |   ||  | ||  |||||     ||||||
Db         83 GTGGGGACAGTGGTNGGTTGCACCTGCCTGCCTGACTATGAGGGTGACGGTTGGAGCTGC 142

Qy       1572 ACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCACCTGT 1631
                ||     |||||||||     || ||||      |||||  ||   |||||||||   |||
Db         143 CGAGCTCGAGACCCCTGCTTGGATGGCCACCGTGGAGGTTGCAGCGAGCACGCTGACTGC 202

Qy       1632 AAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAGATGGG 1691
                |  || || |||  ||  |     |||||  ||  |       |||  ||  ||  ||||||
Db         203 CTCAATACCGGTCCGAACACACGGCGCTGTGAATGCCACGTAGGCTACGTGGGTGATGGG 262

Qy       1692 CTGAACTGTGAGCCGGAGC---AGCTGCCCATTGACCGCTGCTTACAGGACAATGGGCAG 1748
             ||| | |||          |||||   |  |  ||  | |||  | |||||     |       |
Db         263 CTGCAGTGTCTANAGGAGCTTGAACCCCCTGTGGACAGATGCTTGGGAGGATCATCACCT 322

Qy       1749 TGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGTTC 1808
             |||||  | || ||    ||     ||||| || |||||||  |     | ||| || |||
Db         323 TGCCACACTGATGCTTTGTGCACTGACCTACATTTCCAGGAAAAACAGGCTGGTGTCTTC 382

Qy       1809 CATCTACGCTCCCCACTGGGCCAGTAT 1835
             ||  | |   || |    ||||  |||
Db         383 CACATCCAGGCCACCAGTGGCCCTTAT 409


RESULT 3
US-08-024-868-1
; Sequence 1, Application US/08024868
; Patent No. 5386013
;  GENERAL INFORMATION:
;    APPLICANT:  Lee, Tae Ho
;    APPLICANT:  Wisniewski, Hans Georg
;    APPLICANT:  Vilcek, Jan
;    TITLE OF INVENTION:  Cytokine-Induced Protein, TSG-6, DNA
;    TITLE OF INVENTION:  Coding Therefor and Uses Thereof
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  US
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/024,868
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/642,312
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Livnat, Shmuel
;      REGISTRATION NUMBER:  33,949
;      REFERENCE/DOCKET NUMBER:  VILCEK=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-628-5197
;      TELEFAX:  212-737-3528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1414 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      CELL TYPE:  Fibroblast
;      CELL LINE:  FS-4
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  69..899
;      OTHER INFORMATION:
US-08-024-868-1

  Query Match             2.5%;  Score 81.4;  DB 1;  Length 1414;
  Best Local Similarity   55.5%;  Pred. No. 1.3e-13;
  Matches  157;  Conservative    0;  Mismatches  126;  Indels    0;  Gaps    0;

Qy       1795 CTGTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAG 1854
             | |  || |||| |||    |    | |     ||| | || ||||| |||| |   |||
Db        169 CAGCCGGTGTGTACCACAGAGAAGCACGGTCTGGCAAATACAAGCTCACCTACGCAGAAG 228

Qy       1855 CCAGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATG 1914
             | |  | ||  ||||      |||| |      | ||||| ||||| |        |
Db        229 CTAAGGCGGTGTGTGAATTTGAAGGCGGCCATCTCGCAACTTACAAGCAGCTAGAGGCAG 288

Qy       1915 CCCAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCT 1974
              ||   ||       |  ||  | ||  | || || ||| |||  |  ||  | ||||  |
Db        289 CCAGAAAAATTGGATTTCATGTCTGTGCTGCTGGATGGATGGCTAAGGGCAGAGTTGGAT 348

Qy       1975 ACCCCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATG 2034
             ||||||  |      |    | |||||||| | |||      ||| ||  | || ||||
Db        349 ACCCCATTGTGAAGCCAGGGCCCAACTGTGGATTTGGAAAAACTGGCATTATTGATTATG 408

Qy       2035 GACCTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTA 2077
              ||   | | ||| | ||||||||  |||||||  ||  |||||
Db        409 GAATCCGTCTCAATAGGAGTGAAAGATGGGATGCCTATTGCTA 451


RESULT 4
US-08-242-097-1
; Sequence 1, Application US/08242097
; Patent No. 5846763
;  GENERAL INFORMATION:
;    APPLICANT:  Lee, Tae Ho
;    APPLICANT:  Wisniewski, Hans Georg
;    APPLICANT:  Vilcek, Jan
;    TITLE OF INVENTION:  Cytokine-Induced Protein, TSG-6, DNA Coding
;    TITLE OF INVENTION:  Therefor and Uses Thereof
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/242,097
;      FILING DATE:  13-MAY-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  LEE26/VILCEK=1B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-628-5197
;      TELEFAX:  212-737-3528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1414 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      CELL TYPE:  Fibroblast
;      CELL LINE:  FS-4
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  69..899
;      OTHER INFORMATION:
US-08-242-097-1

  Query Match             2.5%;  Score 81.4;  DB 2;  Length 1414;
  Best Local Similarity   55.5%;  Pred. No. 1.3e-13;
  Matches  157;  Conservative    0;  Mismatches  126;  Indels    0;  Gaps    0;

Qy       1795 CTGTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAG 1854
             | |  || |||| |||    |    | |     ||| | || ||||| |||| |   |||
Db        169 CAGCCGGTGTGTACCACAGAGAAGCACGGTCTGGCAAATACAAGCTCACCTACGCAGAAG 228

Qy       1855 CCAGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATG 1914
             | |  | ||  ||||      |||| |      | ||||| ||||| |||||        |
Db        229 CTAAGGCGGTGTGTGAATTTGAAGGCGGCCATCTCGCAACTTACAAGCAGCTAGAGGCAG 288

Qy       1915 CCCAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCT 1974
              ||   ||       |  ||  | ||  | || || ||| |||  |  ||  | ||||  |
Db        289 CCAGAAAAATTGGATTTCATGTCTGTGCTGCTGGATGGATGGCTAAGGGCAGAGTTGGAT 348

Qy       1975 ACCCCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATG 2034
             ||||||  |      |    | |||||||| | |||      ||| ||  | || ||||
Db        349 ACCCCATTGTGAAGCCAGGGCCCAACTGTGGATTTGGAAAAACTGGCATTATTGATTATG 408

Qy       2035 GACCTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTA 2077
              ||   | | ||| | ||||||||  |||||||  ||  |||||
Db        409 GAATCCGTCTCAATAGGAGTGAAAGATGGGATGCCTATTGCTA 451


RESULT 5
US-09-206-695-1
; Sequence 1, Application US/09206695
; Patent No. 6210905
;  GENERAL INFORMATION:
;    APPLICANT:  Lee, Tae Ho
;    APPLICANT:  Wisniewski, Hans Georg
;    APPLICANT:  Vilcek, Jan
;    TITLE OF INVENTION:  Cytokine-Induced Protein, TSG-6, DNA Coding
;    TITLE OF INVENTION:  Therefor and Uses Thereof
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W., Suite 300
```

; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/206,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,097
; FILING DATE: 13-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: LEE26/VILCEK=1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-628-5197
; TELEFAX: 212-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: Fibroblast
; CELL LINE: FS-4
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 69..899
; OTHER INFORMATION:
US-09-206-695-1

Query Match 2.5%; Score 81.4; DB 3; Length 1414;
Best Local Similarity 55.5%; Pred. No. 1.3e-13;
Matches 157; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

```
Qy       1795 CTGTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAG 1854
             | |  || |||| |||   |   | |    ||| | || ||||| |||| |   |||
Db        169 CAGCCGGTGTGTACCACAGAGAAGCACGGTCTGGCAAATACAAGCTCACCTACGCAGAAG 228

Qy       1855 CCAGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATG 1914
             | |  | ||  ||||     ||||  |      | ||||| ||||| |||||       |
Db        229 CTAAGGCGGTGTGTGAATTTGAAGGCGGCCATCTCGCAACTTACAAGCAGCTAGAGGCAG 288

Qy       1915 CCCAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCT 1974
             ||   ||       |  ||  | ||  | || || ||| |||   |  ||  | ||||  |
Db        289 CCAGAAAAATTGGATTTCATGTCTGTGCTGCTGGATGGATGGCTAAGGGCAGAGTTGGAT 348

Qy       1975 ACCCCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATG 2034
             ||||||  |    |   |  |||||||| | |||     ||| ||  | || ||||
Db        349 ACCCCATTGTGAAGCCAGGGCCCAACTGTGGATTTGGAAAAACTGGCATTATTGATTATG 408

Qy       2035 GACCTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTA 2077
             ||    | | ||| | ||||||||  |||||| ||  |||||
Db        409 GAATCCGTCTCAATAGGAGTGAAAGATGGGATGCCTATTGCTA 451
```

RESULT 6
US-09-799-118-1
; Sequence 1, Application US/09799118
; Patent No. 6518401
; GENERAL INFORMATION:
; APPLICANT: Lee, Tae Ho
; Wisniewski, Hans Georg
; Vilcek, Jan
; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
; Therefor and Uses Thereof
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,118
; FILING DATE: 06-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,097
; FILING DATE: 13-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: LEE26/VILCEK=1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-628-5197
; TELEFAX: 212-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: Fibroblast
; CELL LINE: FS-4
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 69..899
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-799-118-1

Query Match 2.5%; Score 81.4; DB 4; Length 1414;
Best Local Similarity 55.5%; Pred. No. 1.3e-13;
Matches 157; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

```
Qy       1795 CTGTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAG 1854
             | |  || |||| |||   |   | |    ||| | || ||||| |||| |   |||
Db        169 CAGCCGGTGTGTACCACAGAGAAGCACGGTCTGGCAAATACAAGCTCACCTACGCAGAAG 228

Qy       1855 CCAGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATG 1914
             | |  | ||  ||||     ||||  |      | ||||| ||||| |||||       |
Db        229 CTAAGGCGGTGTGTGAATTTGAAGGCGGCCATCTCGCAACTTACAAGCAGCTAGAGGCAG 288

Qy       1915 CCCAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCT 1974
             ||   ||       |  ||  | ||  | || || ||| |||   |  ||  | ||||  |
Db        289 CCAGAAAAATTGGATTTCATGTCTGTGCTGCTGGATGGATGGCTAAGGGCAGAGTTGGAT 348

Qy       1975 ACCCCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATG 2034
             ||||||  |    |   |  |||||||| | |||     ||| ||  | || ||||
Db        349 ACCCCATTGTGAAGCCAGGGCCCAACTGTGGATTTGGAAAAACTGGCATTATTGATTATG 408

Qy       2035 GACCTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTA 2077
             ||    | | ||| | ||||||||  |||||| ||  |||||
Db        409 GAATCCGTCTCAATAGGAGTGAAAGATGGGATGCCTATTGCTA 451
```

RESULT 7
US-09-484-970B-63
; Sequence 63, Application US/09484970B

```
; Patent No. 6426186
; GENERAL INFORMATION:
;  APPLICANT:  Jones, Karen A.
;  APPLICANT:  Volkmuth, Wayne
;  APPLICANT:  Walker, Michael G.
;  TITLE OF INVENTION:  BONE REMODELING GENES
;  FILE REFERENCE:  PB-0014 US
;  CURRENT APPLICATION NUMBER: US/09/484,970B
;  CURRENT FILING DATE:  2000-01-18
;  NUMBER OF SEQ ID NOS:  172
;  SOFTWARE:  PERL Program
; SEQ ID NO 63
;   LENGTH: 1734
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6426186 201752.1
;   NAME/KEY: unsure
;   LOCATION: 1444, 1455, 1457, 1461, 1463, 1465, 1467, 1497, 1500, 1506, 1521, 1534-
;   OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-63

  Query Match             2.2%;  Score 72;  DB 4;  Length 1734;
  Best Local Similarity   55.6%;  Pred. No. 9.2e-11;
  Matches  158;  Conservative    0;  Mismatches  125;  Indels    1;  Gaps    1;

Qy       1795 CTGTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAG 1854
             | |  || |||| |||    |    | |      ||| | || ||||| |||| |  |||
Db        169 CAGCCGGTGTGTACCACAGAGAAGCACGGTCTGGCAAATACAAGCTCACCTACGCAGAAG 228

Qy       1855 CCAGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATG 1914
             | |  | ||  ||||    ||||  |      | ||||| ||||| |||||       |
Db        229 CTAAGGCGGTGTGTGAATTTGAAGGCGGCCATCTCGCAACTTACAAGCAGCTAGAGGCAG 288

Qy       1915 CCCAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCT 1974
             ||   ||        |  ||  | ||  | || || ||| |||  |  ||  | ||||  |
Db        289 CCAGAAAAATTGGATTTCATGTCTGTGCTGCTGGATGGATGGCTAAGGGCAGAGTTGGAT 348

Qy       1975 ACCCCACAG-CCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTAT 2033
             ||||||  |    |    ||  |||||||| | |||      ||| ||  | || |||
Db        349 ACCCCATTGTGAAGCCAGGGCCCCAACTGTGGATTTGGAAAAACTGGCATTATTGATTAT 408

Qy       2034 GGACCTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTA 2077
             |||   | | ||| | ||||||||  |||||||  ||  |||||
Db        409 GGAATCCGTCTCAATAGGAGTGAAAGATGGGATGCCTATTGCTA 452


RESULT 8
US-09-907-794A-212
; Sequence 212, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,794A
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
;   LENGTH: 1985
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-907-794A-212

  Query Match             1.8%;  Score 58.6;  DB 4;  Length 1985;
  Best Local Similarity   55.0%;  Pred. No. 9.6e-07;
  Matches  115;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy       1797 GTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCC 1856
             | || ||| || | |    |  |||||    || |  ||  || | | |||  || | | |
Db        593 GGTGTGGTCTTTCCTTACCAGTCCCCCAACGGGCGCTACCAGTTCAACTTCCACGAGGGC 652

Qy       1857 AGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCC 1916
                ||| ||||||  |  | ||||||    ||||  |||   | |||||||  |   |||
Db        653 CAGCAGGTCTGTGCAGAGCAGGCTGCGGTGGTGGCCTCCTTTGAGCAGCTCTTCCGGGCC 712

Qy       1917 CAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTAC 1976
                | |||         ||  ||||   || ||||||||| ||    |     |||    |||
Db        713 TGGGAGGAGGGCCTGGACTGGTGCAACGCGGGCTGGCTGCAGGATGCTACGGTGCAGTAC 772

Qy       1977 CCCACAGCCTTCGCCTCCCAGAACTGTGG 2005
              ||||     ||  ||   |||  ||| ||
Db        773 CCCATCATGTTGCCCCGGCAGCCCTGCGG 801


RESULT 9
```

US-09-905-125A-212
; Sequence 212, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-905-125A-212

Query Match 1.8%; Score 58.6; DB 4; Length 1985;
Best Local Similarity 55.0%; Pred. No. 9.6e-07;
Matches 115; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```
Qy       1797 GTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCC 1856
             | || ||| || | |   |  |||||    || |  ||  || | | |||  || | | |
Db        593 GGTGTGGTCTTTCCTTACCAGTCCCCCAACGGGCGCTACCAGTTCAACTTCCACGAGGGC 652

Qy       1857 AGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCC 1916
              ||| ||||||  |  | ||||||    ||||  |||   | |||||||  |   |||
Db        653 CAGCAGGTCTGTGCAGAGCAGGCTGCGGTGGTGGCCTCCTTTGAGCAGCTCTTCCGGGCC 712

Qy       1917 CAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTAC 1976
               | |||         ||  ||||   || |||||||||  ||   |   |||   |||
Db        713 TGGGAGGAGGGCCTGGACTGGTGCAACGCGGGCTGGCTGCAGGATGCTACGGTGCAGTAC 772

Qy       1977 CCCACAGCCTTCGCCTCCCAGAACTGTGG 2005
              ||||     ||  ||   |||  ||| ||
Db        773 CCCATCATGTTGCCCCGGCAGCCCTGCGG 801
```

RESULT 10
US-09-902-775A-212
; Sequence 212, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547

```
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
;   LENGTH: 1985
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-902-775A-212

  Query Match             1.8%;  Score 58.6;  DB 4;  Length 1985;
  Best Local Similarity   55.0%;  Pred. No. 9.6e-07;
  Matches  115;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy       1797 GTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCC 1856
              | || ||| || | |   | |||||   || |  ||  || | | |||  || | | |
Db        593 GGTGTGGTCTTTCCTTACCAGTCCCCCAACGGGCGCTACCAGTTCAACTTCCACGAGGGC 652

Qy       1857 AGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCC 1916
                ||| ||||||  |  | ||||||   ||||  |||   | |||||||  |   |||
Db        653 CAGCAGGTCTGTGCAGAGCAGGCTGCGGTGGTGGCCTCCTTTGAGCAGCTCTTCCGGGCC 712

Qy       1917 CAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTAC 1976
                | |||          ||  ||||   || |||||||||  ||   |    |||    |||
Db        713 TGGGAGGAGGGCCTGGACTGGTGCAACGCGGGCTGGCTGCAGGATGCTACGGTGCAGTAC 772

Qy       1977 CCCACAGCCTTCGCCTCCCAGAACTGTGG 2005
              ||||     ||  ||   |||  ||| ||
Db        773 CCCATCATGTTGCCCCGGCAGCCCTGCGG 801


RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             1.7%;  Score 54.2;  DB 1;  Length 7218;
  Best Local Similarity   3.6%;  Pred. No. 4.6e-05;
  Matches  14;  Conservative  221;  Mismatches  154;  Indels    0;  Gaps    0;

Qy       1392 AAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACATGCACA 1451
              :::::::: :: : : : ::: : ::   :::  : ::::: :: ::::  : : : : :
Db       1432 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1373

Qy       1452 GTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCTCCCAG 1511
              :  : :::     : ::: :::: :: :::::  : : ::::: ::  ::: :     ::
Db       1372 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1313

Qy       1512 AAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACAGCTGC 1571
              ::::: : :::::      : ::  :  :::::::: : :::::::: ::: : ::  :
Db       1312 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1253

Qy       1572 ACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCACCTGT 1631
              : ::::: :::     : : ::: ::    :: :::::: :  : ::: : :  :   :
Db       1252 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1193

Qy       1632 AAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAGATGGG 1691
              :::: :: :::   ::: ::: : ::: : ::: : ::::  :  : :  ::::: :::
Db       1192 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1133

Qy       1692 CTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGCAGTGC 1751
                :::  : :::  :::: ::  :   :  :: :  :   : :::: :: ::: :: :
Db       1132 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1073

Qy       1752 CATGCAGACGCCAAATGTGTCGACCTCCA 1780
              : : :  ||| |  |   ||||||| ||
Db       1072 RRRRRRATCGCAAGCTCCCTCGACCTGCA 1044


RESULT 12
US-08-001-078A-2
; Sequence 2, Application US/08001078A
; Patent No. 5872094
;  GENERAL INFORMATION:
;    APPLICANT:  Goetinck, Paul F.
;    APPLICANT:  Tondravi, M., Mehrdad
;    APPLICANT:  Binette, Francois
;    TITLE OF INVENTION:  METHODS FOR PROMOTING CARTILAGE MATRIX
;    TITLE OF INVENTION:  FORMATION
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 STATE STREET, Suite 510
;      CITY:  BOSTON
;      STATE:  Massachusetts
;      COUNTRY:  USA
```

```
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  ASCII text
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/001,078A
;       FILING DATE:  06-JAN-1993
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Myers, Paul L.
;       REGISTRATION NUMBER:  35,965
;       REFERENCE/DOCKET NUMBER:  MGP-008
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 227-7400
;       TELEFAX:  (617) 227-5941
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1400 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-001-078A-2

  Query Match             1.7%;  Score 53.8;  DB 2;  Length 1400;
  Best Local Similarity   53.6%;  Pred. No. 2e-05;
  Matches  112;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy       1797 GTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCC 1856
                | || ||| |||| |       | | |||||| |  || || || |  ||| || | ||
Db        706 GGTGTGGTATTCCCTTACTTTCCACGACTGGGGCGCTACAATCTCAATTTTCACGAGGCG 765

Qy       1857 AGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCC 1916
                |||||||     || | | |||     || ||  ||| | ||||||| | | | |||
Db        766 CAGCAGGCCTGTCTGGACCAGGATGCTGTGATCGCCTCCTTCGACCAGCTGTACGACGCC 825

Qy       1917 CAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTAC 1976
                |  || |  |    ||  ||||   || ||||||||        ||    ||     |
Db        826 TGGCGGGGCGGGCTGGACTGGTGCAATGCCGGCTGGCTCAGTGATGGCTCTGTGCAATCT 885

Qy       1977 CCCACAGCCTTCGCCTCCCAGAACTGTGG 2005
                ||||   |      ||    ||  ||||||
Db        886 CCCATCACAAAGCCCAGAGAGCCCTGTGG 914


RESULT 13
US-08-463-218-2
; Sequence 2, Application US/08463218
; Patent No. 5986052
;  GENERAL INFORMATION:
;    APPLICANT:  Goetinck, Paul F.
;    APPLICANT:  Tondravi, Mehrdad
;    APPLICANT:  Binette, Francois
;    TITLE OF INVENTION:  METHODS FOR PROMOTING CARTILAGE MATRIX
;    TITLE OF INVENTION:  FORMATION
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 STATE STREET, Suite 510
;      CITY:  BOSTON
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,218
;      FILING DATE:  05-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/001,078
;      FILING DATE:  06-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Myers, Paul L.
;      REGISTRATION NUMBER:  35,965
;      REFERENCE/DOCKET NUMBER:  MGP-008DV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1400 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-463-218-2

  Query Match             1.7%;  Score 53.8;  DB 2;  Length 1400;
  Best Local Similarity   53.6%;  Pred. No. 2e-05;
  Matches  112;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy       1797 GTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCC 1856
                | || ||| |||| |       | | |||||| |  || || || |  ||| || | ||
Db        706 GGTGTGGTATTCCCTTACTTTCCACGACTGGGGCGCTACAATCTCAATTTTCACGAGGCG 765

Qy       1857 AGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCC 1916
                |||||||     || | | |||     || ||  ||| | ||||||| | | | |||
Db        766 CAGCAGGCCTGTCTGGACCAGGATGCTGTGATCGCCTCCTTCGACCAGCTGTACGACGCC 825

Qy       1917 CAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTAC 1976
                |  || |  |    ||  ||||   || ||||||||        ||    ||     |
Db        826 TGGCGGGGCGGGCTGGACTGGTGCAATGCCGGCTGGCTCAGTGATGGCTCTGTGCAATCT 885

Qy       1977 CCCACAGCCTTCGCCTCCCAGAACTGTGG 2005
                ||||   |      ||    ||  ||||||
Db        886 CCCATCACAAAGCCCAGAGAGCCCTGTGG 914


RESULT 14
PCT-US94-00253-2
; Sequence 2, Application PC/TUS9400253
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  METHODS FOR PROMOTING CARTILAGE MATRIX
;    TITLE OF INVENTION:  FORMATION
;    NUMBER OF SEQUENCES:  2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII (text)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/00253
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/001,078
;      FILING DATE:  06-JAN-1993
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1400 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US94-00253-2

  Query Match             1.7%;  Score 53.8;  DB 5;  Length 1400;
  Best Local Similarity   53.6%;  Pred. No. 2e-05;
  Matches  112;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy       1797 GTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCC 1856
                | || ||| |||| |       | | |||||| |  || || || |  ||| || | ||
Db        706 GGTGTGGTATTCCCTTACTTTCCACGACTGGGGCGCTACAATCTCAATTTTCACGAGGCG 765

Qy       1857 AGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCC 1916
```

```
                 ||||||||    || | | |||    || ||  ||| | ||||||| | | | |||
Db        766 CAGCAGGCCTGTCTGGACCAGGATGCTGTGATCGCCTCCTTCGACCAGCTGTACGACGCC 825

Qy       1917 CAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTAC 1976
              |  || |  |    ||  ||||   || ||||||||       ||    ||    |
Db        826 TGGCGGGGCGGGCTGGACTGGTGCAATGCCGGCTGGCTCAGTGATGGCTCTGTGCAATCT 885

Qy       1977 CCCACAGCCTTCGCCTCCCAGAACTGTGG 2005
              ||||   |     ||    ||  ||||||
Db        886 CCCATCACAAAGCCCAGAGAGCCCTGTGG 914


RESULT 15
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             1.6%;  Score 52.8;  DB 1;  Length 7218;
  Best Local Similarity   2.7%;  Pred. No. 0.00012;
  Matches    9;  Conservative  198;  Mismatches  126;  Indels    0;  Gaps    0;

Qy       2870 TGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAG 2929
              | || ||    ||| | :: :: ::: :: :: :::   :: ::  :: :   :::::
Db       1050 TCGAGGGAGCTTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1109

Qy       2930 CACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTC 2989
              : ::  :: ::::::     : :    ::::::   : :::     :: : :::: ::::
Db       1110 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1169

Qy       2990 TCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTA 3049
              :::  ::  :::     :: ::::: :   :      : : :: :: : ::: :::  :
Db       1170 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1229

Qy       3050 CAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCT 3109
              :  :::::::::::  ::::::  :::::::::::::: : ::::::  ::  : ::: ::
Db       1230 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1289

Qy       3110 CCATTCTGCCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTC 3169
              :: :::: :::: : :  :   :  :: :  ::::::::::: ::   :: :   ::::
Db       1290 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1349

Qy       3170 CNTCTTTGTATCCCAGCCCCTAGCCCAGTGCCT 3202
              : ::::: : ::::  :::::  :::  : :::
Db       1350 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1382


Search completed: June 21, 2004, 16:27:07
Job time : 222 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 10:16:51 ; Search time 7854 Seconds
(without alignments)
12395.048 Million cell updates/sec

Title: US-10-028-248A-210
Perfect score: 3260
Sequence: 1 ggcacgagcaggagcttccc.....aaaaaaaaaaaaaaaaaaaa 3260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1852.6	56.8	3085	11	AK034522	AK034522 Mus muscu
c 2		735.8	22.6	741	14	CA449310	CA449310 UI-H-EI1-
c 3		591	18.1	605	14	CA433190	CA433190 UI-H-CO0-
4		510	15.6	524	13	BX101101	BX101101 BX101101
5		486	14.9	562	12	BG220572	BG220572 RST40358
c 6		472	14.5	481	9	AI081595	AI081595 ox63h05.s
c 7		441.4	13.5	455	9	AI248408	AI248408 qh65e09.x
c 8		418.4	12.8	421	9	AI131339	AI131339 qc18c08.x
9		413.6	12.7	664	10	BB626584	BB626584 BB626584
c 10		404.2	12.4	488	14	R92610	R92610 yq07g08.s1
c 11		401.4	12.3	422	9	AA968560	AA968560 oq75d11.s
12		381	11.7	448	14	H49176	H49176 yq18g02.r1
c 13		378.2	11.6	484	9	AA010567	AA010567 zi09b11.s
14		370.4	11.4	3025	11	BC049247	BC049247 Mus muscu
c 15		358.4	11.0	396	14	H91274	H91274 yu99e07.s1
16		355	10.9	819	13	BX717115	BX717115 BX717115
c 17		333	10.2	334	9	AI084842	AI084842 ow89e11.s
18		332.4	10.2	445	14	H90364	H90364 yu99e07.r1
19		329.8	10.1	459	14	T47504	T47504 yb14f01.r1
c 20		329	10.1	453	14	H49088	H49088 yq18f08.s1
21		328.8	10.1	545	28	AQ278586	AQ278586 CITBI-E1-
c 22		326	10.0	331	14	R97435	R97435 yq53h07.s1
c 23		323	9.9	324	9	AA704436	AA704436 zj22h04.s
c 24		322.6	9.9	414	14	H70480	H70480 yr91e04.s1
25		322.2	9.9	460	14	R97483	R97483 yq53h06.r1
26		314	9.6	325	9	AA257161	AA257161 zr82b05.r
27		313.2	9.6	1155	12	BM919585	BM919585 AGENCOURT
c 28		305.4	9.4	394	9	AA004325	AA004325 zh91f02.s
c 29		301	9.2	302	9	AA257068	AA257068 zr82b05.s
30		300.8	9.2	2368	11	BC019712	BC019712 Mus muscu
31		286.4	8.8	694	14	CA376384	CA376384 654727 NC
32		283.6	8.7	683	12	BJ523552	BJ523552 BJ523552
33		273.2	8.4	348	14	R91674	R91674 yq10f09.r1
c 34		270.8	8.3	328	14	R91675	R91675 yq10f09.s1
35		269.8	8.3	793	12	BI832486	BI832486 603082278
36		247.8	7.6	385	10	BB842118	BB842118 BB842118
37		240	7.4	601	14	CB513007	CB513007 ssalrgb54
c 38		216.4	6.6	228	14	T47505	T47505 yb14f01.s1
39		208.4	6.4	500	9	AA821602	AA821602 vw30c03.r
40		208	6.4	722	13	BX739778	BX739778 BX739778
41		206.8	6.3	462	10	BE656181	BE656181 UI-M-BH0-
42		203	6.2	809	9	AU080852	AU080852 AU080852
43		199	6.1	928	12	BI906024	BI906024 603062355
c 44		194.6	6.0	660	13	BX265613	BX265613 BX265613
45		192	5.9	612	13	BX265614	BX265614 BX265614

ALIGNMENTS

RESULT 1
AK034522
LOCUS AK034522 3085 bp mRNA linear HTC 18-SEP-2003
DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone:9330210I23 product:similar to CD44-LIKE PRECURSOR FELL [Homo sapiens], full insert sequence.
ACCESSION AK034522
VERSION AK034522.1 GI:26330004
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

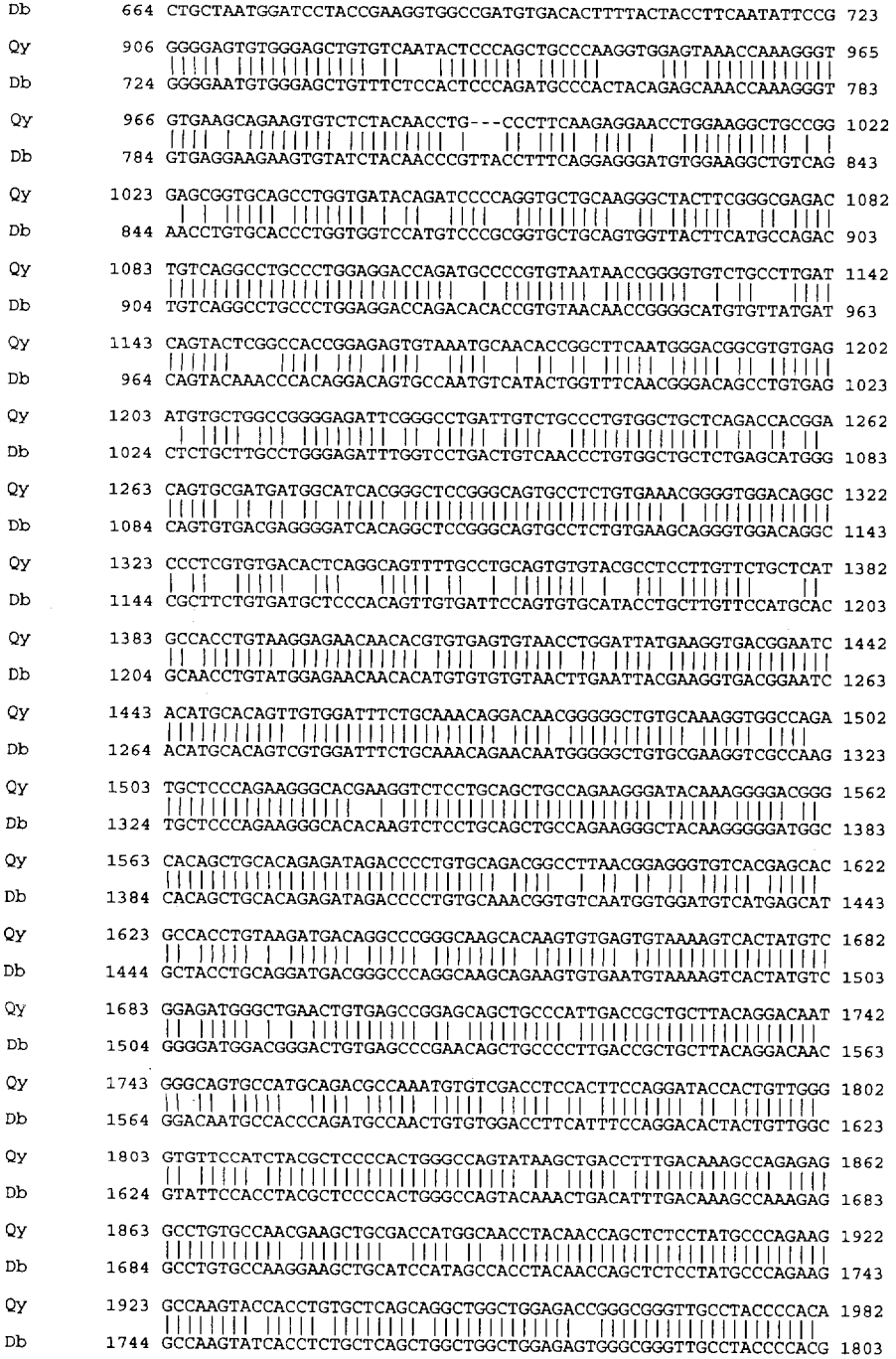
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3085)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers
source 1. .3085
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9330210I23"
/db_xref="MGI:2398728"
/db_xref="taxon:10090"
/clone="9330210I23"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
CDS <1. .2685
/note="unnamed protein product; putative similar to CD44-LIKE PRECURSOR FELL [Homo sapiens] (SPTR|Q9NRY3, evidence: FASTY, 78.1%ID, 99.4%length, match=2673)"
/codon_start=1
/protein_id="BAC28741.1"
/db_xref="GI:26330005"
/translation="GRYHVVACQQLLLENLKVITSATTLQGEPISISVSQDTVLINKK AKVLSSDIISTNGVIHVIDTLLSPQNLLITPKGASGRVLLNLTTVAANHGYTKFSKLI QDSGLLKVITDPMHTPVTLFWPTDKALQALPQEQQDFLFNEDNKDKLKAYLKFHVIRD TMALASDLPRSASWKTLQGSELSVRCGTGSDVGELFLNGQMCRIIQRRLLFDGGVAYG IDCLLMDPTEGGRCDTFTTFNIPGECGSCFSTPRCPLQSKPKGVRKKCIYNPLPFRRD VEGCQNLCTLVVHVPRCCSGYFMPDCQACPGGPDTPCNNRGMCYDQYKPTGQCQCHTG FNGTACELCLPGRFGPDCQPCGCSEHGQCDEGITGSGQCLCEAGWTGRFCDAPTVVIP VCIPACSMHATCMENNTCVCNLNYEGDGITCTVVDFCKQNNGGCAKVAKCSQKGTQVS CSCQKGYKGDGHSCTEIDPCANGVNGGCHEHATCRMTGPGKQKCECKSHYVGDGRDCE PEQLPLDRCLQDNGQCHPDANCVDLHFQDTTVGVFHLRSPLGQYKLTFDKAKEACAKE AASIATYNQLSYAQKAKYHLCSAGWLESGRVAYPTIYASKKCANIVGIVDYGTRTNKS EMWDVFCYRMKDVNCTCKAGYVGDGFSCNGNLLQVLMSFPSLTNFLTEVLVFSRSSAQ GRAFLKHLTDLSISGTLFVPQNSGLPKNKSLSGRDIEHHLTNVNVSFYDDLVNGTVLK TRLGSQLLITSSQDQLHQEARFVDGRAILQWDIIASNGVLHIISEPLKAPPTAATAAH SGLGTGIFCAVVLVTGAIALAAYSYFRLNQRTTGFRRFESEDDIDALAFGKQQPESIT NPLYETSTPAAPEPSCDPFTDSGERELENSDPLGALRS"
polyA_signal 3068. .3073
/note="putative"
polyA_site 3085
/note="putative"

ORIGIN

Query Match 56.8%; Score 1852.6; DB 11; Length 3085;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 2333; Conservative 0; Mismatches 614; Indels 32; Gaps 8;

```
Qy      186 CGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAAT 245
            ||||| || ||||| |||||||| ||||| || ||||| ||||| || ||||| | | |
Db        4 CGGTATCACGTGGTGGCCTGCCAGCAGCTACTGCTGGAGAACCTAAAGGTGATCACGAGT 63

Qy      246 GCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATA 305
            || ||  ||||||||||||||||||||    |||||| |||||||||  ||| |||   |||
Db       64 GCCACAACCCTCCAAGGAGAGCCAATTTCCATCTCTGTCTCTCAGGACACTGTGCTCATA 123

Qy      306 AATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATC 365
            || || ||||| ||| |  | |||||||| || ||||| || || || |  | || ||  ||
Db      124 AACAAGAAGGCGAAGGTTCTGTCCAGTGACATTATCAGCACCAACGGAGTCATCCACGTC 183

Qy      366 ATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGA 425
            ||||||| |||| | |||||| |||| |||||||||| |||||||     ||| || ||
Db      184 ATAGACACATTGTTGTCTCCCCAAAACTTGCTTATCACCCCCAAAGGTGCCTCCGGCAGG 243

Qy      426 ATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTA 485
            |||||| ||| ||||| ||  ||||| ||||| |||| || | |||||| ||||| ||
Db      244 GTTCTGCTAAACCTTACTACAGTGGCAGCAAACCATGGATATACCAAATTCAGCAAGTTG 303

Qy      486 ATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTC 545
            |||||||||||||| |||||||  |||||||| || ||||| || ||||||||||||||||
Db      304 ATACAGGACTCAGGCTTGCTGAAAGTCATCACTGACCCCATGCATACCCCAGTCACTCTC 363

Qy      546 TTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTC 605
            |||||||| || ||| |||||||||| ||| | |||   || || ||||||||||||||||
Db      364 TTCTGGCCTACGGACAAAGCCCTCCAAGCCTTGCCTCAGGAGCAGCAGGACTTCCTGTTC 423

Qy      606 AACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCC 665
            ||  |||||||||| |||||||||||||  ||  |||||| || ||||| ||||| ||
Db      424 AATGAAGACAACAAAGACAAGCTGAAGGCATACCTGAAGTTCCACGTGATCCGAGATACC 483

Qy      666 AAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAG 725
            | || |||||||   || || ||||  ||  || ||||||||||||||||| ||||| |||
Db      484 ATGGCTTTAGCTTCAGACCTCCCCAGGTCTGCTTCCTGGAAGACCCTGCAGGGTTCCGAG 543

Qy      726 CTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACC 785
            |||||||||| |||||| |||||| || |||| || |||||||| ||||| ||||
Db      544 CTGAGTGTGAGGTGTGGAACTGGCAGTGATGTCGGAGAGCTCTTTCTAAATGGACAAATG 603

Qy      786 TGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGT 845
            |||||  ||| | ||||||  |||||||||||  |||||||||| ||||||||||
Db      604 TGCAGGATTATACAGCGGAGACTCTTGTTTGACGGAGGTGTGGCCTATGGCATTGACTGC 663

Qy      846 CTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCG 905
            ||||| || ||||| |||   || ||||| |||||||| |||||||| ||| ||     ||
```


Db 664 CTGCTAATGGATCCTACCGAAGGTGGCCGATGTGACACTTTTACTACCTTCAATATTCCG 723
Qy 906 GGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGT 965
Db 724 GGGGAATGTGGGAGCTGTTTCTCCACTCCCAGATGCCCACTACAGAGCAAACCAAAGGGT 783
Qy 966 GTGAAGCAGAAGTGTCTCTACAACCTG---CCCTTCAAGAGGAACCTGGAAGGCTGCCGG 1022
Db 784 GTGAGGAAGAAGTGTATCTACAACCCGTTACCTTTCAGGAGGGATGTGGAAGGCTGTCAG 843
Qy 1023 GAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGAC 1082
Db 844 AACCTGTGCACCCTGGTGGTCCATGTCCCGCGGTGCTGCAGTGGTTACTTCATGCCAGAC 903
Qy 1083 TGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGAT 1142
Db 904 TGTCAGGCCTGCCCTGGAGGACCAGACACACCGTGTAACAACCGGGGCATGTGTTATGAT 963
Qy 1143 CAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAG 1202
Db 964 CAGTACAAACCCACAGGACAGTGCCAATGTCATACTGGTTTCAACGGGACAGCCTGTGAG 1023
Qy 1203 ATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGA 1262
Db 1024 CTCTGCTTGCCTGGGAGATTTGGTCCTGACTGTCAACCCTGTGGCTGCTCTGAGCATGGG 1083
Qy 1263 CAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGC 1322
Db 1084 CAGTGTGACGAGGGGATCACAGGCTCCGGGCAGTGCCTCTGTGAAGCAGGGTGGACAGGC 1143
Qy 1323 CCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCAT 1382
Db 1144 CGCTTCTGTGATGCTCCCACAGTTGTGATTCCAGTGTGCATACCTGCTTGTTCCATGCAC 1203
Qy 1383 GCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATC 1442
Db 1204 GCAACCTGTATGGAGAACAACACATGTGTGTGTAACTTGAATTACGAAGGTGACGGAATC 1263
Qy 1443 ACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGA 1502
Db 1264 ACATGCACAGTCGTGGATTTCTGCAAACAGAACAATGGGGGCTGTGCGAAGGTCGCCAAG 1323
Qy 1503 TGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGG 1562
Db 1324 TGCTCCCAGAAGGGCACACAAGTCTCCTGCAGCTGCCAGAAGGGCTACAAGGGGGATGGC 1383
Qy 1563 CACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCAC 1622
Db 1384 CACAGCTGCACAGAGATAGACCCCTGTGCAAACGGTGTCAATGGTGGATGTCATGAGCAT 1443
Qy 1623 GCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTC 1682
Db 1444 GCTACCTGCAGGATGACGGGCCCAGGCAAGCAGAAGTGTGAATGTAAAAGTCACTATGTC 1503
Qy 1683 GGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAAT 1742
Db 1504 GGGGATGGACGGGACTGTGAGCCCGAACAGCTGCCCCTTGACCGCTGCTTACAGGACAAC 1563
Qy 1743 GGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGG 1802
Db 1564 GGACAATGCCACCCAGATGCCAACTGTGTGGACCTTCATTTCCAGGACACTACTGTTGGC 1623
Qy 1803 GTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAG 1862
Db 1624 GTATTCCACCTACGCTCCCCACTGGGCCAGTACAAACTGACATTTGACAAAGCCAAAGAG 1683
Qy 1863 GCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAG 1922
Db 1684 GCCTGTGCCAAGGAAGCTGCATCCATAGCCACCTACAACCAGCTCTCCTATGCCCAGAAG 1743
Qy 1923 GCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACA 1982
Db 1744 GCCAAGTATCACCTCTGCTCAGCTGGCTGGCTGGAGAGTGGGCGGGTTGCCTACCCCACG 1803
Qy 1983 GCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGA 2042
Db 1804 ATCTACGCCTCTAAGAAATGTG---CAAACATTGTGGGGATTGTAGACTACGGAACCAGG 1860
Qy 2043 CCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACC 2102
Db 1861 ACCAACAAGAGTGAAATGTGGGACGTCTTCTGTTACAGAATGAAAGATGTTAACTGCACC 1920
Qy 2103 TGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTG 2162
Db 1921 TGCAAGGCAGGCTACGTGGGCGATGGCTTCTCCTGCAATGGGAACCTGCTGCAAGTCCTC 1980
Qy 2163 ATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCA 2222
Db 1981 ATGTCCTTCCCCTCACTCACGAACTTCCTGACAGAGGTGCTGGTTTTTTCCAGGAGCTCA 2040
Qy 2223 GCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTT 2282
Db 2041 GCCCAAGGCCGGGCGTTTTTGAAACACCTGACTGACCTGTCCATCAGTGGCACCCTGTTT 2100
Qy 2283 GTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCAC 2342
Db 2101 GTGCCACAGAACAGTGGGCTACCGAAAAATAAGAGCCTGTCTGGGCGGGACATTGAGCAC 2160
Qy 2343 CACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAA 2402
Db 2161 CACCTCACTAATGTCAACGTCTCCTTTTACGATGACCTTGTCAATGGTACCGTCCTGAAG 2220
Qy 2403 ACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAG 2462
Db 2221 ACTAGGCTGGGAAGCCAACTGCTCATCACCTCCAGCCAGGACCAGCTCCACCA---AGAG 2277
Qy 2463 ACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATC 2522
Db 2278 GCCAGGTTCGTGGATGGAAGAGCCATTCTGCAGTGGGACATCATTGCCTCTAACGGGGTT 2337
Qy 2523 ATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACT 2582
Db 2338 CTCCATATCATTTCTGAACCTTTGAAAGCTCCTCCCACGGCCGCAACGGCTGCCCACTCT 2397
Qy 2583 GGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCT 2642
Db 2398 GGCCTAGGAACAGGCATATTCTGCGCTGTCGTCCTGGTTACTGGTGCGATTGCTCTGGCT 2457
Qy 2643 GCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAA 2702
Db 2458 GCCTACTCCTACTTCCGGCTAAACCAGAGAACAACTGGCTTCCGGCGTTTTGAGTCAGAA 2517
Qy 2703 GAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTG 2762
Db 2518 GATGACATTGATGCCTTGGCCTTTGGCAAGCAGCAGCCTGAGAGTATCACAAACCCTTTG 2577
Qy 2763 TATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAA 2822
Db 2578 TATGAGACCTCAACGCCGGCAGCCCCCGAGCCCTCCTGTGACCCCTTCACCGACTCTGGA 2637
Qy 2823 GAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATG 2882
Db 2638 GAACGGGAGCTGGAGAACAGTGACCCTCTTGGGGCACTG---AGGTCCTGATGTGAGAAG 2694
Qy 2883 CCAGCCATCACTCACTGCCACCTGGGCCATCA----------ACTGTGAATTCTCA 2928
Db 2695 CCAGCCAACAGCCACAGCCACGTGAGCCCTGAGCCATCACGGTTACACGGTGACTCCCCT 2754
Qy 2929 GCACCAGTTGCCTTTTAG---GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCA 2985
Db 2755 GCTCCAGCTGTCACTTGGATCATTTGTTTTAAAGTATGATAACTCAGAAGCCATACCTCA 2814
Qy 2986 TCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCC 3045
Db 2815 CCCCTCTGGTTAGTCTGGGGTTGTCGTCATGGGTAAGGGGACATGTTTCCAGGACACCAG 2874

```
Qy       3046 AGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACC 3105
              |||  || |||||||||  ||||          ||  || ||  ||  | ||| |||  |
Db       2875 TATACCTCTGCCTCCTCTGGGCCTTCACTGTGGCT--CTCTGCACCTTCCAGATGGTGCT 2932

Qy       3106 TGCTCCATTCTGCCCTACATGATGGGTAACTGTGATCTT 3144
              ||  |  ||||||||||  || ||||| |||||| | |
Db       2933 TGGCCTGTTCTGCCCTATCTG-TGGGTAGCTGTGACCCT 2970
```

RESULT 2
CA449310/c

```
LOCUS       CA449310                741 bp    mRNA    linear   EST 08-NOV-2002
DEFINITION  UI-H-EI1-ayt-i-05-0-UI.s1 NCI_CGAP_EI1 Homo sapiens cDNA clone
            UI-H-EI1-ayt-i-05-0-UI 3', mRNA sequence.
ACCESSION   CA449310
VERSION     CA449310.1  GI:24813730
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 20-81, >LINE2 (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-EI1-ayt-i-05-0-UI"
                     /tissue_type="Chondrosarcoma"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_EI1"
                     /note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
                     with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_EI1 is a normalized cDNA library containing the
                     following tissue(s): Chondrosarcoma. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     ACACTTGCAC.
                     TAG_TISSUE=chondrosarcoma
                     TAG_LIB=UI-H-EI1
                     TAG_SEQ=ACACTTGCAC"
ORIGIN
```

```
  Query Match             22.6%;  Score 735.8;  DB 14;  Length 741;
  Best Local Similarity   99.5%;  Pred. No. 2e-124;
  Matches  737;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2516 TGGGATCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGAC 2575
              |||||||||||||||||||||||| ||| ||||||||||||||||||||||||||||||||

Db        741 TGGGATCATTCATGTCATTTCCAGCCCTNTAAAAGCACCCCCTGCCCCCGTGACCTTGAC 682

Qy       2576 CCACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGC 2635
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        681 CCACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTTTTGC 622

Qy       2636 CTTGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGA 2695
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 CTTGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGA 562

Qy       2696 GTCGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAA 2755
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 GTCGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAA 502

Qy       2756 CCCCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGA 2815
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 CCCCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGA 442

Qy       2816 CTCTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACG 2875
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 CTCTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACG 382

Qy       2876 GGAGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAG 2935
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GGAGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAG 322

Qy       2936 TTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGC 2995
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 TTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGC 262

Qy       2996 TGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTT 3055
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 TGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTT 202

Qy       3056 CCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTC 3115
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 CCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTC 142

Qy       3116 TGCCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTT 3175
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        141 TGCCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTT 82

Qy       3176 TGTATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGG 3235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 TGTATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGG 22

Qy       3236 AACAGAAAAAAAAAAAAAAAA 3256
              |||||||||||||||||||||
Db          21 AACAGAAAAAAAAAAAAAAAA 1
```

RESULT 3
CA433190/c

```
LOCUS       CA433190                605 bp    mRNA    linear   EST 08-NOV-2002
DEFINITION  UI-H-CO0-ark-g-11-0-UI.s1 NCI_CGAP_Sub9 Homo sapiens cDNA clone
            UI-H-CO0-ark-g-11-0-UI 3', mRNA sequence.
ACCESSION   CA433190
VERSION     CA433190.1  GI:24797610
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
```

```
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
           from Dr. M. Bento Soares, bento-soares@uiowa.edu
            The following repetitive elements were found in this cDNA
           sequence: 22-B3, >LINE2 (matched compliment)
           Seq primer: M13 FORWARD
           POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-CO0-ark-g-11-0-UI"
                     /tissue_type="mixed"
                     /dev_stage="mixed"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub9"
                     /note="Vector: pT7T3-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoR I; Site_2: Not I; tissues:
                     Cholonic mucosa with Crohns disease, Cholonic mucosa with
                     ulcerative colitis, Fetal thymus, Cervix, Cervical
                     adenosquamous carcinoma, Ligament cells, Prostate
                     carcinoma, Bladder carcinoma, Brain oligodenroga ;
                     NCI_CGAP_Sub9 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are CGTC, AACG, GGGCC,
                     GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional
                     information, contact: Bento Soares, bento-soares@uiowa.edu
                     TAG_TISSUE=colonic mucosa with Crohns disease
                     TAG_LIB=UI-H-CO0
                     TAG_SEQ=CGTC"
ORIGIN

  Query Match             18.1%;  Score 591;  DB 14;  Length 605;
  Best Local Similarity   99.3%;  Pred. No. 6.3e-98;
  Matches  602;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       2653 ACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGACATTA 2712
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 ACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGACATTA 546

Qy       2713 ATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGAGAGCA 2772
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 ATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGAGAGCA 486

Qy       2773 CAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGC 2832
             |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db        485 CAACCTCAGCTCCCCCAGAACCTTCNTACGACCCCTTCACGGACTNTGAAGAACGGCAGC 426

Qy       2833 TTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCA 2892
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCA 366

Qy       2893 CTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGT 2952
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 CTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGT 306

Qy       2953 AAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTT 3012
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 AAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTT 246

Qy       3013 CTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTG 3072
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTG 186

Qy       3073 GCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGT 3132
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 GCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGT 126

Qy       3133 AACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGCCCCTAG 3192
             ||||||| |||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        125 AACTGTGA-CTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGCCCCTAG 67

Qy       3193 CCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAAAAAAAAA 3252
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAAAAAAAA 7

Qy       3253 AAAAAA 3258
             ||||||
Db           6 AAAAAA 1


RESULT 4
BX101101
LOCUS       BX101101                524 bp    mRNA    linear   EST 06-FEB-2003
DEFINITION  BX101101 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGp998O13518 ; IMAGE:241380, mRNA sequence.
ACCESSION   BX101101
VERSION     BX101101.1  GI:27831018
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 524)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998O13518.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .524
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998O13518 ; IMAGE:241380"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             15.6%;  Score 510;  DB 13;  Length 524;
  Best Local Similarity   99.2%;  Pred. No. 4.3e-83;
  Matches  523;  Conservative    0;  Mismatches    1;  Indels    3;  Gaps    1;
```

```
Qy       2729 CAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGAGAGCACAACCTCAGCTCCCCC 2788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGAGAGCACAACCTCAGCTCCCCC 60

Qy       2789 AGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGCTTGAGGGCAATGACCC 2848
             |||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAACCTTCCT---ACCCCTTCACGGACTCTGAAGAACGGCAGCTTGAGGGCAATGACCC 117

Qy       2849 CTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGGG 2908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 CTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGGG 177

Qy       2909 CCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCAC 2968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 CCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCAC 237

Qy       2969 TCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATG 3028
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 TCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATG 297

Qy       3029 TGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGT 3088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 TGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGT 357

Qy       3089 ACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTAACTGTGATCTTTCTT 3148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 ACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTAACTGTGATCTTTCTT 417

Qy       3149 CCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGCCCCTAGCCCAGTGCCTGACACA 3208
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        418 CCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGCCCCTAGCCCAGTGCCTGACACA 477

Qy       3209 GGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAAAAAAAAAAAA 3255
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 GGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAAAAAAAAAAAA 524
```

RESULT 5
BG220572
LOCUS BG220572 562 bp mRNA linear EST 21-APR-2001
DEFINITION RST40358 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG220572
VERSION BG220572.1 GI:13746593
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 562)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 399.
FEATURES Location/Qualifiers
source 1. .562
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."
ORIGIN

Query Match 14.9%; Score 486; DB 12; Length 562;
Best Local Similarity 94.7%; Pred. No. 1e-78;
Matches 534; Conservative 0; Mismatches 27; Indels 3; Gaps 3;

```
Qy       2690 TTTTGAGTCGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATAT 2749
             || ||    |  ||||||||||||||||||| ||||  ||||||| ||||||||||||||
Db          2 TTGTGTCGGGACTGAGGACATTAATGTTGCAGGTCTTACCAAGCAGTAGCCTGAGAATAT 61

Qy       2750 CTCGAACCCCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTT 2809
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CTCGAATCCCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTT 121

Qy       2810 CACGGACTCTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCC 2869
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CACGGACTCTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCC 181

Qy       2870 TGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAG 2929
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        182 TGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCNCAG 241

Qy       2930 CACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTC 2989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTC 301

Qy       2990 TCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTA 3049
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTA 361

Qy       3050 CAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCT 3109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCT 421

Qy       3110 CCATTCTGCCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTC 3169
             |||||||||||||||||||||||||| |||| |||||||||||||||||||| |||
Db        422 CCATTCTGCCCTACATGATGGGTAACTGGGATC-TTCTTCCCTGTTAGATTGTAA-CCTT 479

Qy       3170 CNTCTTTGTATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGT 3229
             | |||||||||||| |||||| |||| |||||||||| | |||||||||  ||| |
Db        480 CAGCTTTGTATCCCA-CCCCTAACCCAATGCCTGACACAAGGACTGTGCACACAAAAAGG 538

Qy       3230 TTATGGAACAGAAAAAAAAAAAAA 3253
             ||||||| ||||||||||||||||
Db        539 TTATGGACCAGAAAAAAAAAAAAA 562
```

RESULT 6
AI081595/c
LOCUS AI081595 481 bp mRNA linear EST 28-AUG-1998
DEFINITION ox63h05.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1661049 3', mRNA sequence.
ACCESSION AI081595
VERSION AI081595.1 GI:3418387
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 481)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

```
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 998   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 460.
FEATURES              Location/Qualifiers
     source           1. .481
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1661049"
                      /tissue_type="Pooled human melanocyte, fetal heart, and
                      pregnant uterus"
                      /lab_host="DH10B"
                      /clone_lib="Soares_NhHMPu_S1"
                      /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                      (Pharmacia) with a modified polylinker; Site_1: Not I;
                      Site_2: Eco RI; Equal amounts of plasmid DNA from three
                      normalized libraries (melanocyte 2NbHM, pregnant uterus
                      NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                      were made in vitro. Following HAP purification, this DNA
                      was used as tracer in a subtractive hybridization
                      reaction. The driver was PCR-amplified cDNAs from pools of
                      5,000 clones made from the same 3 libraries. The pools
                      consisted of I.M.A.G.E. clones 260232-265223,
                      340488-345479, and 484488-489479."
ORIGIN

  Query Match             14.5%;  Score 472;  DB 9;  Length 481;
  Best Local Similarity   98.8%;  Pred. No. 3.9e-76;
  Matches  475;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2774 AACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGCT 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGCT 422

Qy       2834 TGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCAC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCAC 362

Qy       2894 TCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTA 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTA 302

Qy       2954 AAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTC 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTC 242

Qy       3014 TGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGG 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGG 182

Qy       3074 CTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTA 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTA 122

Qy       3134 ACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGCCCCTAGC 3193
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 ACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGCCCCTAGC 62

Qy       3194 CCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAAAAAAAAA 3253
             |||||||||||||||||||||||||||||||||||||||||||||||| |||   ||
Db          61 CCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAAAGTCAAC 2

Qy       3254 A 3254
              |
Db          1 A 1


RESULT 7
AI248408/c
LOCUS       AI248408                455 bp    mRNA    linear   EST 01-DEC-1998
DEFINITION  qh65e09.x1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:1849576 3', mRNA sequence.
ACCESSION   AI248408
VERSION     AI248408.1  GI:3843805
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 455)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 603   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 454.
FEATURES              Location/Qualifiers
     source           1. .455
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1849576"
                      /sex="male"
                      /dev_stage="20 week-post conception fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                      /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                      This is a subtracted version of the original Soares fetal
                      liver spleen 1NFLS library.   1st strand cDNA was primed
                      with a Pac I - oligo(dT) primer [5'
                      AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Pac I and cloned into the Pac I
                      and Eco RI sites of the modified pT7T3 vector.  Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             13.5%;  Score 441.4;  DB 9;  Length 455;
  Best Local Similarity   99.5%;  Pred. No. 1.6e-70;
  Matches  442;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2805 CCCTTCACGGACTCTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGA 2864
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CCCTTCACGGACTCTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGA 385

Qy       2865 GGGCCTGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATT 2924
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 GGGCCTGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATT 325

Qy       2925 CTCAGCACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTC 2984
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 CTCAGCACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTC 265

Qy       2985 ATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACC 3044
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 ATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACC 205

Qy       3045 CAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCAC 3104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 CAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCAC 145

Qy       3105 CTGCTCCATTCTGCCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAA 3164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 CTGCTCCATTCTGCCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAA 85

Qy       3165 GCCTCCNTCTTTGTATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATA 3224
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         84 GCCTCCGTCTTTGTATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATA 25

Qy       3225 AAGGTTTATGGAACAGAAAAAAAA 3248
              |||||||||||||||||| ||||
Db         24 AAGGTTTATGGAACAGAAACAAAA 1


RESULT 8
AI131339/c
LOCUS       AI131339                421 bp    mRNA    linear   EST 27-OCT-1998
DEFINITION  qc18c08.x1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:1709966 3', mRNA sequence.
ACCESSION   AI131339
VERSION     AI131339.1  GI:3601355
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 364   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1709966"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
ORIGIN

  Query Match             12.8%;  Score 418.4;  DB 9;  Length 421;
  Best Local Similarity   99.5%;  Pred. No. 2.5e-66;
  Matches  419;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2824 AACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGC 2883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGC 362

Qy       2884 CAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTT 2943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTT 302

Qy       2944 TAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGG 3003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGG 242

Qy       3004 GGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCT 3063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCT 182

Qy       3064 GACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTAC 3123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTAC 122

Qy       3124 ATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCC 3183
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        121 ATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCC 62

Qy       3184 AGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAA 3243
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          61 AGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCCCAATAAAGGTTTATGGAACAGAAA 2

Qy       3244 A 3244
              |
Db          1 A 1


RESULT 9
BB626584
LOCUS       BB626584                664 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB626584 RIKEN full-length enriched, adult male diencephalon Mus
            musculus cDNA clone 9330210I23 5', mRNA sequence.
ACCESSION   BB626584
VERSION     BB626584.1  GI:16464521
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
            Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
```

```
            e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .664
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="9330210I23"
                     /sex="male"
                     /tissue_type="diencephalon"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, adult male
                     diencephalon"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 185.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC
                     3']. cDNA was cloned into the XhoI and BamHI sites.
                     Vector: a modified pBluescript KS(+) after bulk excision
                     from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end:
                     BamHI"
ORIGIN

  Query Match             12.7%;  Score 413.6;  DB 10;  Length 664;
  Best Local Similarity   77.0%;  Pred. No. 1.8e-65;
  Matches  503;  Conservative    0;  Mismatches  150;  Indels    0;  Gaps    0;

Qy        186 CGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAAT 245
              ||||| || ||||| |||||||| ||||| || ||||| ||||| ||  |||||  | | |
Db          4 CGGTATCACGTGGTGGCCTGCCAGCAGCTACTGCTGGAGAACCTAAAGGTGATCACGAGT 63

Qy        246 GCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATA 305
              || ||  ||||||||||||||||||    |||||| |||||||||  ||| |||   |||
Db         64 GCCACAACCCTCCAAGGAGAGCCAATTTCCATCTCTGTCTCTCAGGACACTGTGCTCATA 123

Qy        306 AATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATC 365
              || || ||||| ||| |  | |||||||| || ||||| || || ||  |  | ||  ||
Db        124 AACAAGAAGGCGAAGGTTCTGTCCAGTGACATTATCAGCACCAACGGAGTCATCCACGTC 183

Qy        366 ATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGA 425
              ||||||| |||| | |||||| |||| |||||||||||| |||||||     ||| || ||
Db        184 ATAGACACATTGTTGTCTCCCCAAAACTTGCTTATCACCCCCAAAGGTGCCTCCGGCAGG 243

Qy        426 ATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTA 485
              |||||| ||| ||||| ||  ||||| ||||| |||| || | |||||| ||||| ||
Db        244 GTTCTGCTAAACCTTACTACAGTGGCAGCAAACCATGGATATACCAAATTCAGCAAGTTG 303

Qy        486 ATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTC 545
              |||||||||||||| |||||||  |||||||| || ||||| || ||||||||||||||||
Db        304 ATACAGGACTCAGGCTTGCTGAAAGTCATCACTGACCCCATGCATACCCCAGTCACTCTC 363

Qy        546 TTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTC 605
              |||||||| || ||| |||||||||| ||| | |||  ||| || ||||||||||||| ||
Db        364 TTCTGGCCTACGGACAAAGCCCTCCAAGCCTTGCCTCATGAGCAGCAGGACTTCCTGCTC 423

Qy        606 AACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCC 665
              ||  |||||||||| ||| |||||||||  ||  |||||||||| ||||| ||| || ||
Db        424 AATGAAGACAACAAAGACCAGCTGAAGGCATACCTGAAGTTTCACGTGATCCGACATACC 483

Qy        666 AAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAG 725
              | || |||||||     | || ||||  ||  |   |||||||||||||| |||||  ||
Db        484 ATGGCTTTAGCTTCAAACCTCCCCAGGTCTGCGTTCTGGAAGACCCTGCATGGTTCCTAG 543
```

```
Qy        726 CTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACC 785
              |||||  || ||| ||||||  ||  |||| || |||||||| ||||| ||||
Db        544 TCTGATGTGACGTGCGGAACTGGCATCGATGTCGGAGAGCTCTTTCTAAATGGACAAATG 603

Qy        786 TGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCAT 838
              ||||  ||| | |||||| | |||||||||||   |||| ||||||  |||||
Db        604 TGCANGATTATACAGCGGAAACTCTTGTTTGACGGAGGTGCGGCCTATGGCAT 656


RESULT 10
R92610/c
LOCUS       R92610                488 bp    mRNA    linear   EST 25-AUG-1995
DEFINITION  yq07g08.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:196286 3', mRNA sequence.
ACCESSION   R92610
VERSION     R92610.1  GI:960150
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 488)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1096
            High quality sequence stops: 267
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1096   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 267.
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3765336"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:196286"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             12.4%;  Score 404.2;  DB 14;  Length 488;
  Best Local Similarity   95.9%;  Pred. No. 9.8e-64;
  Matches  442;  Conservative    0;  Mismatches   16;  Indels    3;  Gaps    3;

Qy       2784 CCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGCTTGAGGGC-AA 2842
              |||| | ||||| |||| |||||||||| |||||||||| || ||||||||||||| ||
Db        463 CCCCNAAAACCTCCCTAGGACCCCTTCAGGGACTCTGAAAAAAGGCAGCTTGAGGGCAAA 404
```

```
Qy       2843 TGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCACTCACTGCCA 2902
             ||||||||||||||| ||||||||| |||||||||||||||| ||||||||||||||||||||
Db        403 TGACCCCTTGAGGANACTGTGAGGNCCTGGACGGGAGATGCNAGCCATCACTCACTGCCA 344

Qy       2903 CCTGGG-CCATCAACTGTGAATTCTCAGCA-CCAGTTGCCTTTTAGGAACGTAAAGTCCT 2960
             || ||| |||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        343 CCCGGGCCCATCAACTGTGAATTCTCAGCACCCAGTTGCCTTTTAGGAACGTAAAGTCCT 284

Qy       2961 TTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGGT 3020
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        283 TTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGANCTGGGGGTTGTTTCTGTGGGT 224

Qy       3021 GAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCT 3080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 GAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCT 164

Qy       3081 TCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTAACTGTGA 3140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 TCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTAACTGTGA 104

Qy       3141 TCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGCCCCTAGCCCAGTGC 3200
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        103 TCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGCCCCTAGCCCAGTGC 44

Qy       3201 CTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGA 3241
             ||||||| |||||||||||||||||||||||||  ||
Db          43 CTGACACNGGAACTGTGCACAATAAAGGTTTATGGAAANGA 3
```

RESULT 11
AA968560/c

LOCUS	AA968560 422 bp mRNA linear EST 17-MAR-1999
DEFINITION	oq75d11.s1 NCI_CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1592181 3', mRNA sequence.
ACCESSION	AA968560
VERSION	AA968560.1 GI:3143740
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 422)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Stratagene, Inc. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Insert Length: 1746 Std Error: 0.00 Seq primer: -40m13 fwd. ET from Amersham High quality sequence stop: 396.

FEATURES	Location/Qualifiers
source	1. .422
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:1592181"
	/sex="mixed"
	/tissue_type="kidney tumor"
	/lab_host="SOLR (kanamycin resistant)"
	/clone_lib="NCI_CGAP_Kid6"
	/note="Organ: kidney; Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled kidney tumors. 5' adaptor sequence: 5' GAATTCGGCACGAG 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."

ORIGIN

Query Match 12.3%; Score 401.4; DB 9; Length 422;
Best Local Similarity 98.6%; Pred. No. 3.3e-63;
Matches 416; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

```
Qy       2840 CAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCACTCACTG 2899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCACTCACTG 363

Qy       2900 CCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTAAAGTCC 2959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTAAAGTCC 303

Qy       2960 TTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGG 3019
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGG 243

Qy       3020 TGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTC 3079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTC 183

Qy       3080 TTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTAACTGTG 3139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTAACTGTG 123

Qy       3140 ATCTTTCTTCCC----TGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGCCCCTAGCCC 3195
             |||||||||||     |||||||||||||||||| ||||||||||||||||||||||||||
Db        122 ATCTTTCTTCCCTGTTTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGCCCCTAGCCC 63

Qy       3196 AGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAAAAAAAAAAAA 3255
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         62 AGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAAAAAAAAAAAAAAAAA 3

Qy       3256 AA 3257
              ||
Db          2 AA 1
```

RESULT 12
H49176

LOCUS	H49176 448 bp mRNA linear EST 22-JAN-1999
DEFINITION	yq18g02.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:274346 5', mRNA sequence.
ACCESSION	H49176
VERSION	H49176.1 GI:989017
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 448)
AUTHORS	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE	The WashU-Merck EST Project
JOURNAL	Unpublished (1995)
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Insert Size: 585 High quality sequence stops: 362 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 585 Std Error: 0.00

```
            Seq primer: M13RP1
            High quality sequence stop: 362.
FEATURES              Location/Qualifiers
     source           1. .448
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="GDB:3798032"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:274346"
                      /sex="male"
                      /dev_stage="20 week-post conception fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares fetal liver spleen 1NFLS"
                      /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                      1st strand cDNA was primed with a Pac I - oligo(dT) primer
                      [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Pac I and cloned into the Pac I
                      and Eco RI sites of the modified pT7T3 vector.  Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match           11.7%;  Score 381;  DB 14;  Length 448;
  Best Local Similarity   99.3%;  Pred. No. 1.8e-59;
  Matches  403;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy       2728 GCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGAGAGCACAACCTCAGCTCCCC 2787
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGAGAGCACAACCTCAGCTCCCC 60

Qy       2788 CAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGCTTGAGGGCAATGACC 2847
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGCTTGAGGGCAATGACC 120

Qy       2848 CCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGG 2907
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCACTCACTGCCACCTGG 180

Qy       2908 GCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCA 2967
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTAAAGTCCTTTAAGCA 240

Qy       2968 CTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGAT 3027
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCTGTGGGTGAGAGAT 300

Qy       3028 GTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTG 3087
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGCTCTTCTTCCTTTG 360

Qy       3088 TACTCTTCAGCT-GGCACCTGCTCCATTC-TGCCCTACATGATGGG 3131
             |||||||||||| |||||||||||||||  |||||||||||||||
Db        361 TACTCTTCAGCTGGGCACCTGCTCCATTNTTGCCCTACATGATGGG 406


RESULT 13
AA010567/c
LOCUS       AA010567                484 bp    mRNA    linear   EST 09-MAY-1997
DEFINITION  zi09b11.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:430269 3', mRNA sequence.
ACCESSION   AA010567
VERSION     AA010567.1  GI:1471613
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 484)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 946   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 307.
FEATURES              Location/Qualifiers
     source           1. .484
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="GDB:1330038"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:430269"
                      /sex="male"
                      /dev_stage="20 week-post conception fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                      /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                      This is a subtracted version of the original Soares fetal
                      liver spleen 1NFLS library.  1st strand cDNA was primed
                      with a Pac I - oligo(dT) primer [5'
                      AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Pac I and cloned into the Pac I
                      and Eco RI sites of the modified pT7T3 vector.  Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match           11.6%;  Score 378.2;  DB 9;  Length 484;
  Best Local Similarity   97.2%;  Pred. No. 5.6e-59;
  Matches  413;  Conservative    0;  Mismatches    8;  Indels    4;  Gaps    3;

Qy       2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880
             |||||||||| ||||||| |||||||||||||||||||| ||||||||||||||||||||
Db        425 AAGAACGGCANCTTGAGGNCAATGACCCCTTGAGGACACTNTGAGGGCCTGGACGGGAGA 366

Qy       2881 TGCCAGCCATCACTCACTGCCA--CCTGGGCCATCAACTGTGAATTCTCAGCACC-AGTT 2937
             |||||||||||||||||||||  |  || ||||||||||||||||||||||| ||||
Db        365 TGCCAGCCATCACTCACTGCCAACCTGGGCCCATCAACTGTGAATTCTCAGCACCNAGTT 306

Qy       2938 GCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 2997
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GCCTTTTAGGAACGTTAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 246

Qy       2998 ATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCC 3057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCC 186

Qy       3058 TCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTC-AGCTGGCACCTGCTCCATTCT 3116
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        185 TCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAAGCTGGCACCTGCTCCATTCT 126

Qy       3117 GCCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTT 3176
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        125 GCCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTT 66

Qy       3177 GTATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGA 3236
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         65 GTATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGA 6

Qy       3237 ACAGA 3241

Db          5 ACAGA 1
```

RESULT 14
BC049247

```
LOCUS       BC049247                3025 bp    mRNA    linear   HTC 28-MAR-2003
DEFINITION  Mus musculus, Similar to stabilin 1, clone IMAGE:5038209, mRNA.
ACCESSION   BC049247
VERSION     BC049247.1  GI:29351595
KEYWORDS    HTC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3025)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 99 Row: g Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..3025
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="CZECH II"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5038209"
                     /tissue_type="Mammary tumor metastatized to lung. Tumor
                     arose spontaneously from a senescent normal mammary
                     (clonal) outgrowth infected with the virus MMTV."
                     /clone_lib="NCI_CGAP_Lu29"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

  Query Match             11.4%;  Score 370.4;  DB 11;  Length 3025;
  Best Local Similarity   54.8%;  Pred. No. 1.1e-57;
  Matches  828;  Conservative    0;  Mismatches  666;  Indels   18;  Gaps    4;

Qy       1046 GATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTCAGGCCTGCCCTGGAGGACC 1105

Db       1020 GAAGCCCAGCTGTTGCCCTGGTCACTATGGCATTAACTGCCATGCTTGCCCTGGTGGCCC 1079

Qy       1106 AGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGTACTCGGCCACCGGAGAGTG 1165

Db       1080 CAGAAGCCCCTGCAGTGACCATGGCGTGTGCCTGGATGGTATACGTGGAAGTGGGCAGTG 1139

Qy       1166 TAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGTGCTGGCCGGGGAGATTCGG 1225

Db       1140 TAATTGCCATCCGGGTTTTGCTGGGACAGCCTGTGAACTCTGTGCCCCAGGTGCCTTTGG 1199

Qy       1226 GCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGTGCGATGATGGCATCACGGG 1285

Db       1200 ACCCCAGTGCCAAGCTTGCCGCTGTACCCAACATGGCCGTTGTGATGAAGGTCTTGGGGG 1259

Qy       1286 CTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCTCGTGTGACACTCAGGCAGT 1345

Db       1260 CTCTGGCTCCTGCTTCTGCGATGAAGGCTGGACTGGGGCACGCTGTGAAGTGCAGCTGGA 1319

Qy       1346 TTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCACCTGTAAGGAGAACAACAC 1405

Db       1320 GTTGCAGCCTGTGTGCACCCCACCGTGTGCACCCCAGGCCGTGTGCCGTCTGGGCAACAG 1379

Qy       1406 GTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACATGCACAGTTGTGGATTTCTG 1465

Db       1380 CTGTGAGTGCAGCCTGGGCTATGAAGGGGATGGCCGTGTGTGCACAGTGGCAGACCTGTG 1439

Qy       1466 CAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCTCCCAGAAGGGCACGAAGGT 1525

Db       1440 CCAGAAGGGGCATGGTGGCTGCAGTAAGCATGCCAACTGCAGCCAGGTGGGGACAGTGGT 1499

Qy       1526 CTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACAGCTGCACAGAGATAGACCC 1585

Db       1500 CACTTGCACCTGCCTGCCTGACTATGAGGGTGACGGTTGGAGCTGCCGAGCTCGAGACCC 1559

Qy       1586 CTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCACCTGTAAGATGACAGGCCC 1645

Db       1560 CTGCTTGGATGGCCACCGTGGAGGTTGCAGCGAGCACGCTGACTGCCTCAATACCGGTCC 1619

Qy       1646 GGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAGATGGGCTGAACTGTGAGCC 1705

Db       1620 GAACACACGGCGCTGTGAATGCCACGTAGGCTACGTGGGTGATGGGCTGCAGTGTCTAGA 1679

Qy       1706 GGAGC---AGCTGCCCATTGACCGCTGCTTACAGGACAATGGGCAGTGCCATGCAGACGC 1762

Db       1680 GGAGCTTGAACCCCCTGTGGACAGATGCTTGGGAGGATCATCACCTTGCCACACTGATGC 1739

Qy       1763 CAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGTTCCATCTACGCTCCCC 1822

Db       1740 TTTGTGCACTGACCTACATTTCCAGGAAAAACAGGCTGGTGTCTTCCACATCCAGGCCAC 1799

Qy       1823 ACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCTGTGCCAACGAAGCTGC 1882

Db       1800 CAGTGGCCCTTATGGTCTGACCTTCTCAGAGGCCACGGAAGCGTGTGAGGGCCAGGGAGC 1859

Qy       1883 GACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAAGTACCACCTGTGCTC 1942

Db       1860 CGTCCTTGCTTCACTCCCTCAACTCTCTGCCGCCCAACAGCTAGGTTTTCATGTCTGCTT 1919

Qy       1943 AGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCTTCGCCTCCCAGAACTG 2002

Db       1920 CGTGGGCTGGTTGGCCAACGGCTCTGCTGCCCACCCTGTTGTCACCCCAGCGGCAGACTG 1979

Qy       2003 TGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCAACAAGAGTGAAATGTG 2062

Db       1980 TGGTAATAATCGTGTAGGTGTAGTTAGCCTTGGGGTCCGTAAGAACCTCTCGGAGCTCTG 2039

Qy       2063 GGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCAAGGTGGGCTATGTGGG 2122

Db       2040 GGATGCCTACTGTTACCGCGTGCAAGACGTGGCTTGCCAGTGTCGGGCCGGCTTCGTGGG 2099

Qy       2123 AGATGG---CTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGTCCTTCCCCTCACT 2179

Db       2100 TGACGGGATCAGCACGTGCAACGGGAAACTGCTTGATGTCTTGGCCGCCACTGCCAACTT 2159
```

```
Qy       2180 CACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTCGAGGCCGTGCATT 2239
             | | | ||||    |    |||||| |||| ||||   | |     |||| | || |||
Db       2160 CTCCACCTTCTATGGGATGCTGCTGGGCTATGCCAATGCCACCCAGAGAGGTCTTGAATT 2219

Qy       2240 TCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCCACAGAACAGTGG 2299
             | | ||   ||||   ||   |  || |  |   || ||||| || ||    ||||  ||
Db       2220 TATGGATTTCCTGGAGGATGAGCTCACCTACAAGACACTCTTCGTTCCTGTCAACAAAGG 2279

Qy       2300 GCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACCTCGCCAATGTCAG 2359
               | |||| ||   |||  ||  ||| |  ||  | || | ||||    | |||| ||
Db       2280 CTTTGTGGACAACATGACGCTGAGTGGCCCAGATCTAGAACTCCACGCTTCTAATGCCAC 2339

Qy       2360 CATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGAGGGTGGGAAGCAA 2419
             | |  |     ||||| |     |   || ||    || |   |      ||     |
Db       2340 CTTTCTGAGTGTCAATGCC---AGCCGGGGGACATTGCTTCCTGCCCACTCAGGTCTTAG 2396

Qy       2420 GCTGCTCATCACTGCCAG---------CCAGGACCCACTCCAACCGACGGAGACCAGGTT 2470
              ||  |||| | || ||           | |     | ||    ||   ||   |  ||
Db       2397 CCTCTTCATAAGTGACATGGGCCCTGACAACACTTCTCTGGTTCCTCTGGCCCCGGGGGC 2456

Qy       2471 TGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCATGT 2530
              ||||  |  ||   | |  |  ||||||||| | || | ||| || ||||||||||
Db       2457 AGTTGTGGTTAGCCACGTCATCGTGTGGGACATCATGGCTTTCAACGGCATCATTCATGC 2516

Qy       2531 CATTTCCAGGCC 2542
                |  |||| ||
Db       2517 TCTGGCCAGCCC 2528
```

RESULT 15
H91274/c

```
LOCUS       H91274                 396 bp    mRNA    linear   EST 28-NOV-1995
DEFINITION  yu99e07.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:241380 3', mRNA sequence.
ACCESSION   H91274
VERSION     H91274.1  GI:1081704
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 360
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: Promega -21m13
            High quality sequence stop: 360.
FEATURES             Location/Qualifiers
     source          1. .396
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3790513"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:241380"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN
```

```
  Query Match             11.0%;  Score 358.4;  DB 14;  Length 396;
  Best Local Similarity   99.0%;  Pred. No. 2.4e-55;
  Matches  381;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    2;

Qy       2837 GGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATG-CCAGCCATCACTC 2895
             || |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        385 GGCCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCCAGCCATCACTC 326

Qy       2896 ACTGCCACCTGGG-CCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTAA 2954
             |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        325 ACTGCCACCTGGGCCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGTAA 266

Qy       2955 AGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCT 3014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 AGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTTCT 206

Qy       3015 GTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGC 3074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTGGC 146

Qy       3075 TCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTAA 3134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 TCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGTAA 86

Qy       3135 CTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGCCCCTAGCC 3194
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         85 CTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGCCCCTAGCC 26

Qy       3195 CAGTGCCTGACACAGGAACTGTGCA 3219
             |||||||||||||||||||||||||
Db         25 CAGTGCCTGACACAGGAACTGTGCA 1
```

Search completed: June 21, 2004, 16:23:23
Job time : 7864 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 10:14:56 ; Search time 1177 Seconds
(without alignments)
11766.462 Million cell updates/sec

Title: US-10-028-248A-210
Perfect score: 3260
Sequence: 1 ggcacgagcaggagcttccc.....aaaaaaaaaaaaaaaaaaaa 3260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3259	100.0	3260	6	ABT08489	Abt08489 Human nov
2	3232.6	99.2	4576	6	ABA04662	Aba04662 Human Hya
3	3232.6	99.2	4642	7	ABX13822	Abx13822 cDNA enco
4	2999.8	92.0	8444	6	ABT08487	Abt08487 Human nov
5	2995.6	91.9	8495	6	ABT08488	Abt08488 Human nov
6	2853.2	87.5	4962	7	ABX13823	Abx13823 cDNA enco
7	2828.8	86.8	3625	4	AAF87120	Aaf87120 NOV9 codi
8	1936.8	59.4	4706	6	ABA04648	Aba04648 Rat Hyalu
9	1936.8	59.4	4706	7	ABX13821	Abx13821 cDNA enco
10	1778.6	54.6	2483	4	AAF87119	Aaf87119 NOV8 codi
11	1777	54.5	2512	4	AAF87117	Aaf87117 NOV6 codi
12	1721.2	52.8	2011	4	AAF87113	Aaf87113 NOV2 codi
13	1429.6	43.9	1804	4	AAF87114	Aaf87114 NOV3 codi
14	1347.4	41.3	1377	3	AAC76373	Aac76373 Human ORF
15	1227.4	37.7	1259	3	AAA57365	Aaa57365 DNA encod
16	990.2	30.4	1624	4	AAF87118	Aaf87118 NOV7 codi
17	764.2	23.4	1577	4	AAF87121	Aaf87121 NOV10 cod
18	708	21.7	718	5	AAS00854	Aas00854 Human cDN
19	537	16.5	722	6	AAS62772	Aas62772 cDNA sequ
20	424.6	13.0	3681	6	ABL90758	Abl90758 Human pol
21	423.4	13.0	7879	8	ACF04000	Acf04000 Human CLE
22	421.8	12.9	6761	3	AAA57362	Aaa57362 DNA encod
23	408.8	12.5	448	8	ACH43442	Ach43442 Human foe
24	391.2	12.0	2863	9	ADB62456	Adb62456 Human cDN
25	285	8.7	2160	9	ADB63670	Adb63670 Human cDN
26	259.6	8.0	2753	9	ADB47730	Adb47730 Novel hum
27	248	7.6	1522	3	AAA57363	Aaa57363 DNA encod
28	247.4	7.6	395	8	ACH29524	Ach29524 Human adu
29	187.4	5.7	1482	7	ACA55632	Aca55632 Human sig
30	143	4.4	473	8	ACH15198	Ach15198 Human adu
31	135	4.1	420	8	ACH21247	Ach21247 Human adu
32	125.2	3.8	474	8	ACH45565	Ach45565 Human foe
33	81.4	2.5	781	7	ABX74436	Abx74436 Human cDN
34	81.4	2.5	1144	7	ABX63446	Abx63446 Human cDN
35	81.4	2.5	1144	9	ADE25720	Ade25720 Human cDN
36	81.4	2.5	1414	2	AAQ27190	Aaq27190 Tumour ne
37	81.4	2.5	1414	2	AAV71778	Aav71778 Tumour ne
38	81.4	2.5	1414	4	AAD06019	Aad06019 Human tum
39	81.4	2.5	1414	6	ABS54635	Abs54635 Human cDN
40	81.4	2.5	1422	7	ACC72678	Acc72678 Human can
41	81.4	2.5	1422	7	ABX76382	Abx76382 Lung canc
42	81.4	2.5	1728	7	ABX63445	Abx63445 Human cDN
c 43	78.2	2.4	351	6	ABV95472	Abv95472 Human pan
44	78.2	2.4	1414	6	ABL68341	Abl68341 Kidney ca
45	78.2	2.4	1414	6	ABK83989	Abk83989 Human cDN

ALIGNMENTS

RESULT 1
ABT08489
ID ABT08489 standard; cDNA; 3260 BP.
XX
AC ABT08489;
XX
DT 28-NOV-2002 (first entry)
XX
DE Human novel protein NOV1c coding sequence SEQ ID NO: 210.
XX
KW Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
KW antiaddictive; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(887,C)
FT /*tag= a
FT variation replace(1034,T)
FT /*tag= b
FT variation replace(1223,T)
FT /*tag= c
XX
PN WO200259315-A2.
XX
PD 01-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050076.
XX
PR 19-DEC-2000; 2000US-0256619P.
PR 19-JAN-2001; 2001US-0262959P.
PR 28-FEB-2001; 2001US-0272408P.
PR 20-APR-2001; 2001US-0285189P.
PR 26-JUL-2001; 2001US-0308039P.
PR 09-AUG-2001; 2001US-0311266P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;
PI Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;
PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;

PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
PI Rothenberg M;
XX
DR WPI; 2002-666903/71.
DR P-PSDB; ABJ10588.
XX
PT New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders e.g.
PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
PT or Alzheimer's disease.
XX
PS Claim 42; Page 24-25; 363pp; English.
XX
CC The present invention provides the protein and coding sequences of
CC several novel human proteins, designated NOVX. These can be used in the
CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
CC Lindau syndrome, Alzheimer's disease, stroke, tuberous sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
CC graft-versus-host disease. The present sequence is a coding sequence of
CC the invention
XX
SQ Sequence 3260 BP; 774 A; 886 C; 850 G; 749 T; 0 U; 1 Other;

Query Match 100.0%; Score 3259; DB 6; Length 3260;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGG 60

Qy         61 AGCATTTCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCATTTCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTG 120

Qy        121 CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGG 180

Qy        181 TTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCT 240

Qy        241 CAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGT 300

Qy        301 ATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTC 360

Qy        361 ATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTG 420

Qy        421 GAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCA 480

Qy        481 ACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCA 540

Qy        541 CTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600

Qy        601 TGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAG 660

Qy        661 ATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTT 720

Qy        721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCC 780

Qy        781 AAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTG 840

Qy        841 ACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG 900

Qy        901 CCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAA 960

Qy        961 AGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCC 1020

Qy       1021 GGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAG 1080

Qy       1081 ACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTG 1140

Qy       1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG 1200

Qy       1201 AGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACG 1260

Qy       1261 GACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAG 1320

Qy       1321 GCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTC 1380

Qy       1381 ATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAA 1440

Qy       1441 TCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCA 1500

Qy       1501 GATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACG 1560

Qy       1561 GGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGC 1620

Qy       1621 ACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATG 1680

Qy       1681 TCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACA 1740
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 TCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACA 1740

Qy        1741 ATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTG 1800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 ATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTG 1800

Qy        1801 GGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 GGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860

Qy        1861 AGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGA 1920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 AGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGA 1920

Qy        1921 AGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCA 1980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1921 AGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCA 1980

Qy        1981 CAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTA 2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 CAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTA 2040

Qy        2041 GACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCA 2100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2041 GACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCA 2100

Qy        2101 CCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCC 2160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 CCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCC 2160

Qy        2161 TGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCT 2220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 TGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCT 2220

Qy        2221 CAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCT 2280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 CAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCT 2280

Qy        2281 TTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGC 2340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 TTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGC 2340

Qy        2341 ACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGC 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 ACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGC 2400

Qy        2401 AAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGG 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 AAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGG 2460

Qy        2461 AGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGA 2520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 AGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGA 2520

Qy        2521 TCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACA 2580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 TCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACA 2580

Qy        2581 CTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGG 2640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 CTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGG 2640

Qy        2641 CTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGG 2700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 CTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGG 2700

Qy        2701 AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCT 2760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCT 2760

Qy        2761 TGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTG 2820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2761 TGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTG 2820

Qy        2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2821 AAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGA 2880

Qy        2881 TGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCC 2940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2881 TGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCC 2940

Qy        2941 TTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATC 3000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2941 TTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATC 3000

Qy        3001 TGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCC 3060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3001 TGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCC 3060

Qy        3061 TCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCC 3120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3061 TCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCC 3120

Qy        3121 TACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTAT 3180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3121 TACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTAT 3180

Qy        3181 CCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3181 CCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAG 3240

Qy        3241 AAAAAAAAAAAAAAAAAAAA 3260
               ||||||||||||||||||||
Db        3241 AAAAAAAAAAAAAAAAAAAA 3260
```

```
RESULT 2
ABA04662
ID   ABA04662 standard; cDNA; 4576 BP.
XX
AC   ABA04662;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Human Hyaluronic Acid Receptor for Endocytosis, HARE, coding sequence.
XX
KW   HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;
KW   chondroitin sulphate; extracellular matrix; cartilage; skin;
KW   vitreous humour; endocytic receptor; glycosaminoglycan; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .4185
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Human HARE"
FT                   /note= "No start codon given"
XX
PN   WO200181544-A2.
XX
PD   01-NOV-2001.
XX
PF   25-APR-2001; 2001WO-US013403.
XX
PR   25-APR-2000; 2000US-0199538P.
PR   02-NOV-2000; 2000US-0245320P.
XX
PA   (WEIG/) WEIGEL P A.
PA   (ZHOU/) ZHOU B.
PA   (WEIG/) WEIGEL J A.
XX
PI   Weigel PA,  Zhou B,  Weigel JA;
XX
```

DR WPI; 2002-049271/06.
DR P-PSDB; AAM47684.
XX
PT New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for
PT identifying agents that inhibit binding to hyaluronic acid, and related
PT nucleic acid.
XX
PS Claim 49; Fig 33; 263pp; English.
XX
CC The present invention relates to sequences for rat and human HARE
CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675
CC and AAM47684). HARE can bind specifically to at least one of hyaluronic
CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin
CC sulphate (CS). HA is an extracellular matrix component of all tissues, in
CC particular cartilage, skin and vitreous humour. HARE is the endocytic
CC receptor responsible for removing HA and other glycosaminoglycans from
CC the circulation. The present sequence is the coding sequence for human
CC HARE
XX
SQ Sequence 4576 BP; 1127 A; 1193 C; 1203 G; 1053 T; 0 U; 0 Other;

Query Match 99.2%; Score 3232.6; DB 6; Length 4576;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3238; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 1387

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 1447

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1448 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 1507

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1508 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 1567

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1568 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 1627

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1628 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 1687

Qy        367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1688 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 1747

Qy        427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1748 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 1807

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1808 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 1867

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1868 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 1927

Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1928 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 1987

Qy        667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1988 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 2047

Qy        727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2048 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 2107

Qy        787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2108 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 2167

Qy        847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2168 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 2227

Qy        907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2228 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 2287

Qy        967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2288 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 2347

Qy       1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2348 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 2407

Qy       1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2408 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 2467

Qy       1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2468 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 2527

Qy       1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
              |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db       2528 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 2587

Qy       1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2588 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 2647

Qy       1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2648 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 2707

Qy       1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2708 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 2767

Qy       1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2768 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 2827

Qy       1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2828 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 2887

Qy       1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2888 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 2947

Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2948 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 3007

Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3008 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 3067

Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3068 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 3127

Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
```

```
Db       3128 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 3187
Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
Db       3188 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 3247
Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
Db       3248 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 3307
Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
Db       3308 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 3367
Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
Db       3368 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 3427
Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
Db       3428 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 3487
Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
Db       3488 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 3547
Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
Db       3548 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 3607
Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
Db       3608 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 3667
Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
Db       3668 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 3727
Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
Db       3728 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 3787
Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
Db       3788 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 3847
Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
Db       3848 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 3907
Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
Db       3908 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 3967
Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
Db       3968 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 4027
Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
Db       4028 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 4087
Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
Db       4088 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 4147
Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
Db       4148 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 4207
Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
Db       4208 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 4267
Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
Db       4268 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 4327
Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
Db       4328 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 4387
Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126
Db       4388 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 4447
Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186
Db       4448 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 4507
Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246
Db       4508 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 4567
Qy       3247 AAAAAAAA 3254
Db       4568 AGTCAACA 4575
```

RESULT 3
ABX13822
ID ABX13822 standard; cDNA; 4642 BP.
XX
AC ABX13822;
XX
DT 19-FEB-2003 (first entry)
XX
DE cDNA encoding human 190kDa Hyaluronan receptor for endocytosis #1.
XX
KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
KW tumour; gene therapy; human; gene; ss.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	CDS	1. .4251
FT		/*tag= a
FT		/product= "HARE"
FT		/note= "Hyaluronan receptor for endocytosis"
FT		/partial
FT		/note= "No start codon given"

XX
PN WO200286093-A2.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013209.
XX
PR 25-APR-2001; 2001US-00842930.
PR 25-APR-2001; 2001US-0286468P.
XX
PA (WEIG/) WEIGEL P H.
PA (WEIG/) WEIGEL J A.
XX
PI Weigel PH, Weigel JA;
XX
DR WPI; 2003-093126/08.
DR P-PSDB; ABG72499.
XX
PT Targeting compounds e.g. chemotherapeutic agent to cell of subject
PT expressing functional active hyaluronan receptor for endocytosis of HARE,
PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
PT epitope.
XX

PS Example; Fig 9A; 167pp; English.
XX
CC The invention describes a method of targeting a compound to a cell or
CC tissue of an individual expressing a functionally active hyaluronan (HA)
CC receptor for endocytosis (HARE) or a cell that does not express
CC functionally active HARE. The method involves using HA molecule, a
CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
CC a monoclonal antibody raised against a HA-binding domain of HARE. The
CC method is useful for targeting a compound, preferably a chemotherapeutic
CC agent or a radioisotope to cell of an individual, especially a human,
CC expressing HARE on its surface (e.g. gene therapy). Also described is a
CC method useful for preventing interaction between a cell having at least
CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
CC its surface. This second method is useful for preventing metastasis by
CC preventing interaction between tumour cells having HA, CD or CDS coat and
CC non-tumour cells expressing HARE on its surface. The invention also
CC describes a method useful for detecting the presence of HA, CD and CDS in
CC a biological fluid. This sequence encodes the human 190kDa Hyaluronan
CC receptor for endocytosis (HARE)
XX
SQ Sequence 4642 BP; 1141 A; 1215 C; 1216 G; 1070 T; 0 U; 0 Other;

Query Match 99.2%; Score 3232.6; DB 7; Length 4642;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3238; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy       7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
Db    1394 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 1453

Qy      67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
Db    1454 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 1513

Qy     127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
Db    1514 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 1573

Qy     187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
Db    1574 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 1633

Qy     247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
Db    1634 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 1693

Qy     307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
Db    1694 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 1753

Qy     367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
Db    1754 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 1813

Qy     427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
Db    1814 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 1873

Qy     487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
Db    1874 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 1933

Qy     547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
Db    1934 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 1993

Qy     607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
Db    1994 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 2053

Qy     667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
Db    2054 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 2113

Qy     727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
Db    2114 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 2173

Qy     787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
Db    2174 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 2233

Qy     847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
Db    2234 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 2293

Qy     907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
Db    2294 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 2353

Qy     967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
Db    2354 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 2413

Qy    1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
Db    2414 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 2473

Qy    1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
Db    2474 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 2533

Qy    1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
Db    2534 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 2593

Qy    1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
Db    2594 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 2653

Qy    1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
Db    2654 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 2713

Qy    1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
Db    2714 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 2773

Qy    1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
Db    2774 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 2833

Qy    1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
Db    2834 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 2893

Qy    1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
Db    2894 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 2953

Qy    1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
Db    2954 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 3013

Qy    1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
Db    3014 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 3073

Qy    1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
Db    3074 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 3133

Qy    1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
Db    3134 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 3193
```

```
Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3194 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 3253

Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3254 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 3313

Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3314 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 3373

Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3374 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 3433

Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3434 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 3493

Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3494 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 3553

Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3554 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 3613

Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3614 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 3673

Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3674 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 3733

Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3734 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 3793

Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3794 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 3853

Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3854 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 3913

Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3914 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 3973

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3974 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 4033

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4034 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 4093

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4094 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 4153

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4154 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 4213

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4214 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 4273

Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4274 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 4333

Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4334 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 4393

Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4394 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 4453

Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4454 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 4513

Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       4514 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 4573

Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       4574 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 4633

Qy       3247 AAAAAAAA 3254
             |  ||| |
Db       4634 AGTCAACA 4641
```

```
RESULT 4
ABT08487
ID   ABT08487 standard; cDNA; 8444 BP.
XX
AC   ABT08487;
XX
DT   28-NOV-2002  (first entry)
XX
DE   Human novel protein NOV1a coding sequence SEQ ID NO: 1.
XX
KW   Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
KW   antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
KW   haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
KW   antidepressant; immunosuppressive; antibacterial; antiparasitic;
KW   virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
KW   antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
KW   antiaddictive; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(4185,C)
FT                   /*tag=  a
FT   variation       replace(4524,C)
FT                   /*tag=  b
FT   variation       replace(4654,A)
FT                   /*tag=  c
FT   variation       replace(4671,G)
FT                   /*tag=  d
FT   variation       replace(4820,C)
FT                   /*tag=  e
XX
PN   WO200259315-A2.
XX
PD   01-AUG-2002.
XX
PF   19-DEC-2001; 2001WO-US050076.
XX
PR   19-DEC-2000; 2000US-0256619P.
PR   19-JAN-2001; 2001US-0262959P.
PR   28-FEB-2001; 2001US-0272408P.
PR   20-APR-2001; 2001US-0285189P.
PR   26-JUL-2001; 2001US-0308039P.
PR   09-AUG-2001; 2001US-0311266P.
XX
```

PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;
PI Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;
PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
PI Rothenberg M;
XX
DR WPI; 2002-666903/71.
DR P-PSDB; ABJ10586.
XX
PT New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders e.g.
PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
PT or Alzheimer's disease.
XX
PS Claim 8; Page 17-19; 363pp; English.
XX
CC The present invention provides the protein and coding sequences of
CC several novel human proteins, designated NOVX. These can be used in the
CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
CC Lindau syndrome, Alzheimer's disease, stroke, tuberous sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
CC graft-versus-host disease. The present sequence is a coding sequence of
CC the invention
XX
SQ Sequence 8444 BP; 2251 A; 2105 C; 2147 G; 1929 T; 0 U; 12 Other;

Query Match 92.0%; Score 2999.8; DB 6; Length 8444;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 3128; Conservative 0; Mismatches 18; Indels 108; Gaps 1;

```
Qy       7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5279 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5338

Qy      67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5339 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 5398

Qy     127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5399 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 5458

Qy     187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5459 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 5518

Qy     247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5519 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 5578

Qy     307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5579 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 5638

Qy     367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5639 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 5698

Qy     427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5699 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 5758

Qy     487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5759 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 5818

Qy     547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5819 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 5878

Qy     607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5879 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 5938

Qy     667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5939 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 5998

Qy     727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5999 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 6058

Qy     787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6059 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 6118

Qy     847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6119 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 6178

Qy     907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6179 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 6238

Qy     967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6239 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 6298

Qy    1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6299 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 6358

Qy    1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6359 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 6418

Qy    1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6419 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 6478

Qy    1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6479 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 6538

Qy    1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6539 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 6598

Qy    1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6599 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 6658

Qy    1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6659 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 6718

Qy    1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6719 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 6778

Qy    1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6779 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 6838

Qy    1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6839 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 6898
```

```
Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6899 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 6958

Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6959 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 7018

Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7019 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 7078

Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7079 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 7138

Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7139 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 7198

Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7199 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 7258

Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7259 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 7318

Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
             ||||||||||||||||||||||||||||||||||||||||
Db       7319 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAA---------------------- 7358

Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7359 ---------------------------------------------------------------- 7358

Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
                                                 ||||||||||||||||||||||||||||||
Db       7359 ----------------------------GGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 7390

Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7391 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 7450

Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7451 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 7510

Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7511 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 7570

Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
             || ||||||||||||||||||||||||||||||||||||||||||||||||    | | |
Db       7571 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGGTACAAAGTA 7630

Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7631 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 7690

Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7691 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 7750

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
             ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       7751 TGGGAGCAGGGATCTTCTTTTGCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 7810

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       7811 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTACCAGCATTTTGAGTCGGAAGAGG 7870

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7871 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 7930

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7931 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 7990

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7991 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAG 8050

Qy       2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 2946
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8051 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAG 8110

Qy       2947 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 3006
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8111 GAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGG 8170

Qy       3007 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 3066
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8171 TTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGAC 8230

Qy       3067 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 3126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8231 CCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATG 8290

Qy       3127 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGC 3186
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       8291 ATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGC 8350

Qy       3187 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAA 3246
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       8351 CCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAA 8410

Qy       3247 AAAAAAAAAAAAAA 3260
                |   ||||||||||
Db       8411 AGTCAAAAAAAAAA 8424
```

```
RESULT 5
ABT08488
ID   ABT08488 standard; cDNA; 8495 BP.
XX
AC   ABT08488;
XX
DT   28-NOV-2002  (first entry)
XX
DE   Human novel protein NOV1b coding sequence SEQ ID NO: 3.
XX
KW   Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
KW   antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
KW   haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
KW   antidepressant; immunosuppressive; antibacterial; antiparasitic;
KW   virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
KW   antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
KW   antiaddictive; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200259315-A2.
XX
PD   01-AUG-2002.
XX
PF   19-DEC-2001; 2001WO-US050076.
XX
PR   19-DEC-2000; 2000US-0256619P.
PR   19-JAN-2001; 2001US-0262959P.
PR   28-FEB-2001; 2001US-0272408P.
PR   20-APR-2001; 2001US-0285189P.
PR   26-JUL-2001; 2001US-0308039P.
PR   09-AUG-2001; 2001US-0311266P.
XX
```

PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;
PI Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;
PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
PI Rothenberg M;
XX
DR WPI; 2002-666903/71.
DR P-PSDB; ABJ10587.
XX
PT New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders e.g.
PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
PT or Alzheimer's disease.
XX
PS Claim 8; Page 21-23; 363pp; English.
XX
CC The present invention provides the protein and coding sequences of
CC several novel human proteins, designated NOVX. These can be used in the
CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
CC Lindau syndrome, Alzheimer's disease, stroke, tuberous sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
CC graft-versus-host disease. The present sequence is a coding sequence of
CC the invention
XX
SQ Sequence 8495 BP; 2295 A; 2102 C; 2135 G; 1957 T; 0 U; 6 Other;

Query Match 91.9%; Score 2995.6; DB 6; Length 8495;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 3137; Conservative 0; Mismatches 10; Indels 116; Gaps 2;

```
Qy       7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5311 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 5370

Qy      67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5371 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 5430

Qy     127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5431 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 5490

Qy     187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5491 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 5550

Qy     247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5551 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 5610

Qy     307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5611 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 5670

Qy     367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5671 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 5730

Qy     427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5731 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 5790

Qy     487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5791 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 5850

Qy     547 TCTGGCCCACCGACCAA---------GCCCTCCATGCCCTACCTGCTGAACAACAGGACT 597
          |||||||||||||||||         |||||||||||||||||||||||||||||||||||||
Db    5851 TCTGGCCCACCGACCAAGCCCTCCATGCCCTCCATGCCCTACCTGCTGAACAACAGGACT 5910

Qy     598 TCCTGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATAC 657
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5911 TCCTGTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATAC 5970

Qy     658 GAGATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAG 717
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5971 GAGATGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAG 6030

Qy     718 GTTCAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATG 777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6031 GTTCAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATG 6090

Qy     778 GCCAAACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCA 837
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6091 GCCAAACCTACAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCA 6150

Qy     838 TTGACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCG 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6151 TTGACTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCG 6210

Qy     898 ATGCCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAAC 957
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6211 ATGCCTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAAC 6270

Qy     958 CAAAGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCT 1017
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6271 CAAAGGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCT 6330

Qy    1018 GCCGGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGC 1077
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6331 GCCGGGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGC 6390

Qy    1078 GAGACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCC 1137
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6391 GAGACTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCC 6450

Qy    1138 TTGATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGT 1197
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6451 TTGATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGT 6510

Qy    1198 GTGAGATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACC 1257
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6511 GTGAGATGTGCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACC 6570

Qy    1258 ACGGACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGA 1317
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6571 ACGGACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGA 6630

Qy    1318 CAGGCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTG 1377
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6631 CAGGCCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTG 6690

Qy    1378 CTCATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACG 1437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6691 CTCATGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACG 6750

Qy    1438 GAATCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGG 1497
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6751 GAATCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGG 6810

Qy    1498 CCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGG 1557
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6811 CCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGG 6870

Qy    1558 ACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACG 1617
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6871 ACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACG 6930
```

```
Qy       1618 AGCACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACT 1677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6931 AGCACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACT 6990

Qy       1678 ATGTCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGG 1737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6991 ATGTCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGG 7050

Qy       1738 ACAATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTG 1797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7051 ACAATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTG 7110

Qy       1798 TTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCA 1857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7111 TTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCA 7170

Qy       1858 GAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCC 1917
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7171 GAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCC 7230

Qy       1918 AGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACC 1977
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7231 AGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACC 7290

Qy       1978 CCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGAC 2037
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7291 CCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGAC 7350

Qy       2038 CTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACT 2097
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7351 CTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAA---------- 7400

Qy       2098 GCACCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGG 2157
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7401 ---------------------------------------------------------- 7400

Qy       2158 TCCTGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACA 2217
             ||||||||||||||||||||||||||
Db       7401 ---------------------------GGAAGTGCTGGCCTATTCCAACA 7423

Qy       2218 GCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCC 2277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7424 GCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCC 7483

Qy       2278 TCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCG 2337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7484 TCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCG 7543

Qy       2338 AGCACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCC 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7544 AGCACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCC 7603

Qy       2398 TGCAAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGA 2457
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7604 TGCAAACGAGGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGA 7663

Qy       2458 CGGAGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATG 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7664 CGGAGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATG 7723

Qy       2518 GGATCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCC 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7724 GGATCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCC 7783

Qy       2578 ACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCT 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7784 ACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCT 7843

Qy       2638 TGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGT 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7844 TGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGT 7903

Qy       2698 CGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACC 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7904 CGGAAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACC 7963

Qy       2758 CCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACT 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7964 CCTTGTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACT 8023

Qy       2818 CTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGG 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8024 CTGAAGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGG 8083

Qy       2878 AGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTT 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8084 AGATGCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTT 8143

Qy       2938 GCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 2997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8144 GCCTTTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 8203

Qy       2998 ATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCC 3057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8204 ATCTGGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCC 8263

Qy       3058 TCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTG 3117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8264 TCCTCTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTG 8323

Qy       3118 CCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTG 3177
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       8324 CCCTACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTG 8383

Qy       3178 TATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAA 3237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8384 TATCCCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAA 8443

Qy       3238 CAGAAAAAAAAAAAAAAAAAAA 3260
             |||||| |||   ||||||||||
Db       8444 CAGAAACAAAGTCAAAAAAAAAA 8466
```

```
RESULT 6
ABX13823
ID   ABX13823 standard; cDNA; 4962 BP.
XX
AC   ABX13823;
XX
DT   19-FEB-2003  (first entry)
XX
DE   cDNA encoding human 190kDa Hyaluronan receptor for endocytosis #2.
XX
KW   Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
KW   chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
KW   tumour; gene therapy; human; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .4962
FT                   /*tag= a
FT                   /product= "HARE"
FT                   /note= "Hyaluronan receptor for endocytosis"
FT                   /partial
FT                   /note= "No start codon given"
XX
PN   WO200286093-A2.
XX
PD   31-OCT-2002.
XX
PF   25-APR-2002; 2002WO-US013209.
XX
PR   25-APR-2001; 2001US-00842930.
PR   25-APR-2001; 2001US-0286468P.
XX
```

PA (WEIG/) WEIGEL P H.
PA (WEIG/) WEIGEL J A.
XX
PI Weigel PH, Weigel JA;
XX
DR WPI; 2003-093126/08.
DR P-PSDB; ABG72514.
XX
PT Targeting compounds e.g. chemotherapeutic agent to cell of subject
PT expressing functional active hyaluronan receptor for endocytosis of HARE,
PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
PT epitope.
XX
PS Example; Fig 9B; 167pp; English.
XX
CC The invention describes a method of targeting a compound to a cell or
CC tissue of an individual expressing a functionally active hyaluronan (HA)
CC receptor for endocytosis (HARE) or a cell that does not express
CC functionally active HARE. The method involves using HA molecule, a
CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
CC a monoclonal antibody raised against a HA-binding domain of HARE. The
CC method is useful for targeting a compound, preferably a chemotherapeutic
CC agent or a radioisotope to cell of an individual, especially a human,
CC expressing HARE on its surface (e.g. gene therapy). Also described is a
CC method useful for preventing interaction between a cell having at least
CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
CC its surface. This second method is useful for preventing metastasis by
CC preventing interaction between tumour cells having HA, CD or CDS coat and
CC non-tumour cells expressing HARE on its surface. The invention also
CC describes a method useful for detecting the presence of HA, CD and CDS in
CC a biological fluid. This sequence encodes a longer version of the human
CC 190kDa Hyaluronan receptor for endocytosis (HARE) shown in ABG72499
CC (Encoded by ABX13822)
XX
SQ Sequence 4962 BP; 1249 A; 1277 C; 1311 G; 1125 T; 0 U; 0 Other;

Query Match 87.5%; Score 2853.2; DB 7; Length 4962;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2855; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 66
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2105 ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATT 2164

Qy         67 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2165 TCGTGAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCT 2224

Qy        127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2225 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 2284

Qy        187 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2285 GGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATG 2344

Qy        247 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2345 CTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAA 2404

Qy        307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 366
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2405 ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCA 2464

Qy        367 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2465 TAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAA 2524

Qy        427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2525 TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAA 2584

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2585 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 2644

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2645 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 2704

Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2705 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 2764

Qy        667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2765 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 2824

Qy        727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2825 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 2884

Qy        787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2885 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 2944

Qy        847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2945 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 3004

Qy        907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3005 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 3064

Qy        967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3065 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 3124

Qy       1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3125 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 3184

Qy       1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3185 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 3244

Qy       1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3245 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 3304

Qy       1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3305 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 3364

Qy       1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3365 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 3424

Qy       1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3425 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 3484

Qy       1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3485 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 3544

Qy       1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3545 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 3604

Qy       1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3605 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 3664

Qy       1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3665 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 3724

Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3725 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 3784

Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3785 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 3844

Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3845 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 3904

Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3905 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 3964

Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3965 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 4024

Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4025 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 4084

Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4085 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 4144

Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4145 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 4204

Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4205 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 4264

Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4265 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 4324

Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4325 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 4384

Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4385 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 4444

Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4445 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 4504

Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 2466
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4505 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCA 4564

Qy       2467 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 2526
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4565 GGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTC 4624

Qy       2527 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4625 ATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCT 4684

Qy       2587 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 2646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4685 TGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTT 4744

Qy       2647 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 2706
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4745 ACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGG 4804

Qy       2707 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 2766
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4805 ACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATG 4864

Qy       2767 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 2826
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4865 AGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAAC 4924

Qy       2827 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGA 2864
             ||||||||||||||||||||||||||||||||||||||
Db       4925 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGA 4962
```

RESULT 7
AAF87120
ID AAF87120 standard; DNA; 3625 BP.
XX
AC AAF87120;
XX
DT 26-MAR-2002 (first entry)
XX
DE NOV9 coding sequence.
XX
KW NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW cardiovascular; casein kinase II phosphorylation site; contraception;
KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW epidermal growth factor; cell development; apoptosis; cell adhesion;
KW growth migration; cell structure; motility; cancer; immune disorder;
KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.
XX
OS Unidentified.
XX
FH Key   Location/Qualifiers
FT CDS   387. .2024
FT   /*tag= a
FT   /product= "NOV9"
XX
PN WO200136638-A2.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-US031543.
XX
PR 19-NOV-1999; 99US-0166336P.
PR 29-NOV-1999; 99US-0167785P.
PR 08-MAR-2000; 2000US-0187844P.
PR 16-NOV-2000; 2000US-00715417.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Lichenstein H, Vernet C, Fernandes E;
XX
DR WPI; 2001-648134/74.
DR P-PSDB; AAB83365.
XX
PT Novel human polypeptides and the nucleic acids that encode them useful
PT for preventing, diagnosing and treating e.g. cancer, inflammation and
PT immune disorders.
XX
PS Claim 9; Page 32-33; 141pp; English.
XX
CC This sequence encodes the NOV9 protein. The invention relates to the NOV1
CC -NOV16 proteins, and their coding sequences. The proteins have Cytostatic
CC ; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
CC activities. The sequences may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate NOVX expression. They
CC may be used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC activity of protein by expressing inactive proteins or to supplement the

CC patients own production of protein. They are used to produce NOVX
CC proteins, by inserting the nucleic acid into a cell and culturing it to
CC express the protein. The DNA may be used as DNA probes in assays to
CC detect and quantitate the presence of similar DNAs in samples, and which
CC patients may need restorative therapy. The NOVX protein may also be used
CC as antigens in the production of antibodies (Abs) against NOVX and in
CC assays to identify modulators of NOVX expression and activity. The anti-
CC NOVX Abs and antagonist are used to down regulate expression and
CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX
CC in samples. Disorders that may be prevented, diagnosed and/or treated
CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
CC NOV13-16 have casein kinase II phosphorylation sites characteristic of
CC serine/threonine kinases, and are used to treat kinase-related disorders
CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
CC -like super family and are involved in, e.g. regulation of cell
CC development, apoptosis, cell adhesion, growth migration, cell structure
CC and motility and protein management, and are used to treat cancers,
CC inflammatory disorders, immune disorders and cellular adhesion disorders.
CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
CC syndrome and marfan syndrome
XX
SQ Sequence 3625 BP; 896 A; 971 C; 914 G; 844 T; 0 U; 0 Other;

Query Match 86.8%; Score 2828.8; DB 4; Length 3625;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2985; Conservative 0; Mismatches 43; Indels 111; Gaps 2;

```
Qy       122 AGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGT 181
Db       569 ACCCTCTGATCTTCTGTTACAGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGT 628

Qy       182 TCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTC 241
Db       629 TCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTC 688

Qy       242 AAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTA 301
Db       689 AAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTA 748

Qy       302 TATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCA 361
Db       749 TATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCA 808

Qy       362 TATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGG 421
Db       809 TATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGG 868

Qy       422 AAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAA 481
Db       869 AAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAA 928

Qy       482 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCAC 541
Db       929 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCAC 988

Qy       542 TCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCT 601
Db       989 TCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCT 1048

Qy       602 GTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 661
Db       1049 GTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 1108

Qy       662 TGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTC 721
Db       1109 TGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTC 1168

Qy       722 AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCA 781
Db       1169 AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCA 1228
```

```
Qy       782 AACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGA 841
Db       1229 AACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGA 1288

Qy       842 CTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGC 901
Db       1289 CTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGC 1348

Qy       902 CTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAA 961
Db       1349 CTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAA 1408

Qy       962 GGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCG 1021
Db       1409 GGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCG 1468

Qy       1022 GGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGA 1081
Db       1469 GGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGA 1528

Qy       1082 CTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGA 1141
Db       1529 CTGTCAG----GGTGAGGGTGCCTCTTCCCCCCTCGCAACTCTAAAAGTGTCTGCCTTGA 1584

Qy       1142 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA 1201
Db       1585 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA 1644

Qy       1202 GATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGG 1261
Db       1645 GATGTGCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGG 1704

Qy       1262 ACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGG 1321
Db       1705 ACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGG 1764

Qy       1322 CCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCA 1381
Db       1765 CCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCA 1824

Qy       1382 TGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAAT 1441
Db       1825 TGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAAT 1884

Qy       1442 CACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAG 1501
Db       1885 CACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAG 1944

Qy       1502 ATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGG 1561
Db       1945 ATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGG 2004

Qy       1562 GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCA 1621
Db       2005 GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCA 2064

Qy       1622 CGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGT 1681
Db       2065 CGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGT 2124

Qy       1682 CGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAA 1741
Db       2125 CGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAA 2184

Qy       1742 TGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGG 1801
Db       2185 TGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGG 2244

Qy       1802 GGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA 1861
Db       2245 GGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA 2304

Qy       1862 GGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAA 1921
```

```
Db       2305 GGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAA 2364
Qy       1922 GGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCAC 1981
Db       2365 GGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCAC 2424
Qy       1982 AGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAG 2041
Db       2425 AGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAG 2484
Qy       2042 ACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCAC 2101
Db       2485 ACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAA---------------- 2530
Qy       2102 CTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCT 2161
Db       2531 ------------------------------------------------------------ 2530
Qy       2162 GATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTC 2221
Db       2531 --------------------------------GGAAGTGCTGGCCTATTCCAACAGCTC 2557
Qy       2222 AGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTT 2281
Db       2558 AGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTT 2617
Qy       2282 TGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCA 2341
Db       2618 TGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCA 2677
Qy       2342 CCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCA 2401
Db       2678 CCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCA 2737
Qy       2402 AACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGA 2461
Db       2738 AACGAGGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGA 2797
Qy       2462 GACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGAT 2521
Db       2798 GACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGAT 2857
Qy       2522 CATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACAC 2581
Db       2858 CATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACAC 2917
Qy       2582 TGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGC 2641
Db       2918 TGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGC 2977
Qy       2642 TGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGA 2701
Db       2978 TGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGA 3037
Qy       2702 AGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTT 2761
Db       3038 AGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTT 3097
Qy       2762 GTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGA 2821
Db       3098 GTATGAGAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGA 3157
Qy       2822 AGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGAT 2881
Db       3158 AGAACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGAT 3217
Qy       2882 GCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCT 2941
Db       3218 GCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCT 3277
Qy       2942 TTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCT 3001
Db       3278 TTTAGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCT 3337
Qy       3002 GGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCT 3061
Db       3338 GGGGGTTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCT 3397
Qy       3062 CTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCT 3121
Db       3398 CTGACCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCT 3457
Qy       3122 ACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATC 3181
Db       3458 ACATGATGGGTAACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATC 3517
Qy       3182 CCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGA 3241
Db       3518 CCAGCCCCTAGCCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGA 3577
Qy       3242 AAAAAAAAAAAAAAAAAAA 3260
Db       3578 AACAAAGTCAAAAAAAAAA 3596
```

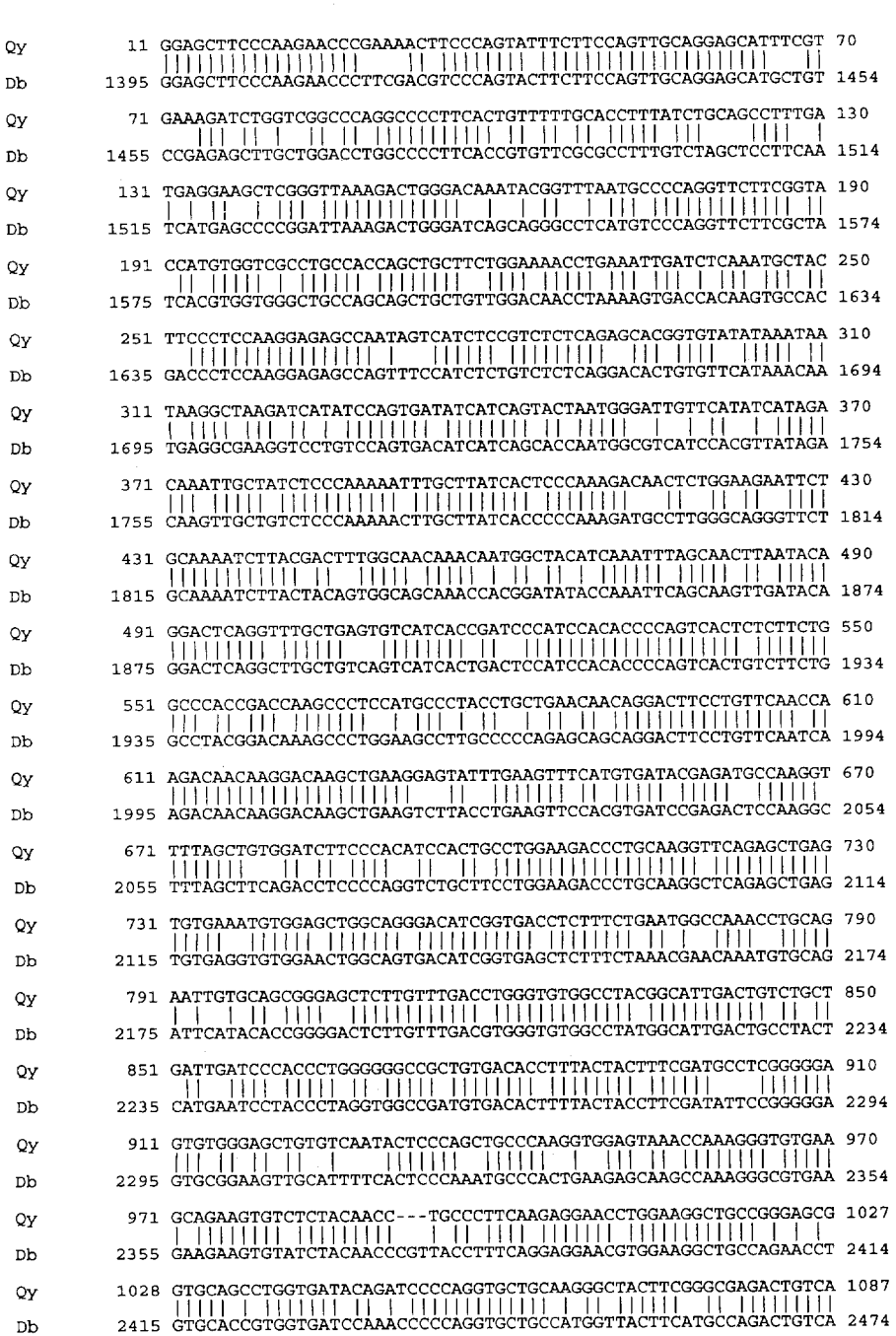
RESULT 8
ABA04648
ID ABA04648 standard; cDNA; 4706 BP.
XX
AC ABA04648;
XX
DT 22-FEB-2002 (first entry)
XX
DE Rat Hyaluronic Acid Receptor for Endocytosis, HARE, coding sequence.
XX
KW HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;
KW chondroitin sulphate; extracellular matrix; cartilage; skin;
KW vitreous humour; endocytic receptor; glycosaminoglycan; rat; ss.
XX
OS Rattus norvegicus.
XX
FH Key   Location/Qualifiers
FT CDS   1. .4296
FT   /*tag= a
FT   /partial
FT   /product= "Rat HARE"
FT   /note= "No start codon given"
XX
PN WO200181544-A2.
XX
PD 01-NOV-2001.
XX
PF 25-APR-2001; 2001WO-US013403.
XX
PR 25-APR-2000; 2000US-0199538P.
PR 02-NOV-2000; 2000US-0245320P.
XX
PA (WEIG/) WEIGEL P A.
PA (ZHOU/) ZHOU B.
PA (WEIG/) WEIGEL J A.
XX
PI Weigel PA, Zhou B, Weigel JA;
XX
DR WPI; 2002-049271/06.
DR P-PSDB; AAM47675.
XX
PT New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for
PT identifying agents that inhibit binding to hyaluronic acid, and related
PT nucleic acid.
XX
PS Claim 49; Fig 21; 263pp; English.
XX
CC The present invention relates to sequences for rat and human HARE
CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675
CC and AAM47684). HARE can bind specifically to at least one of hyaluronic

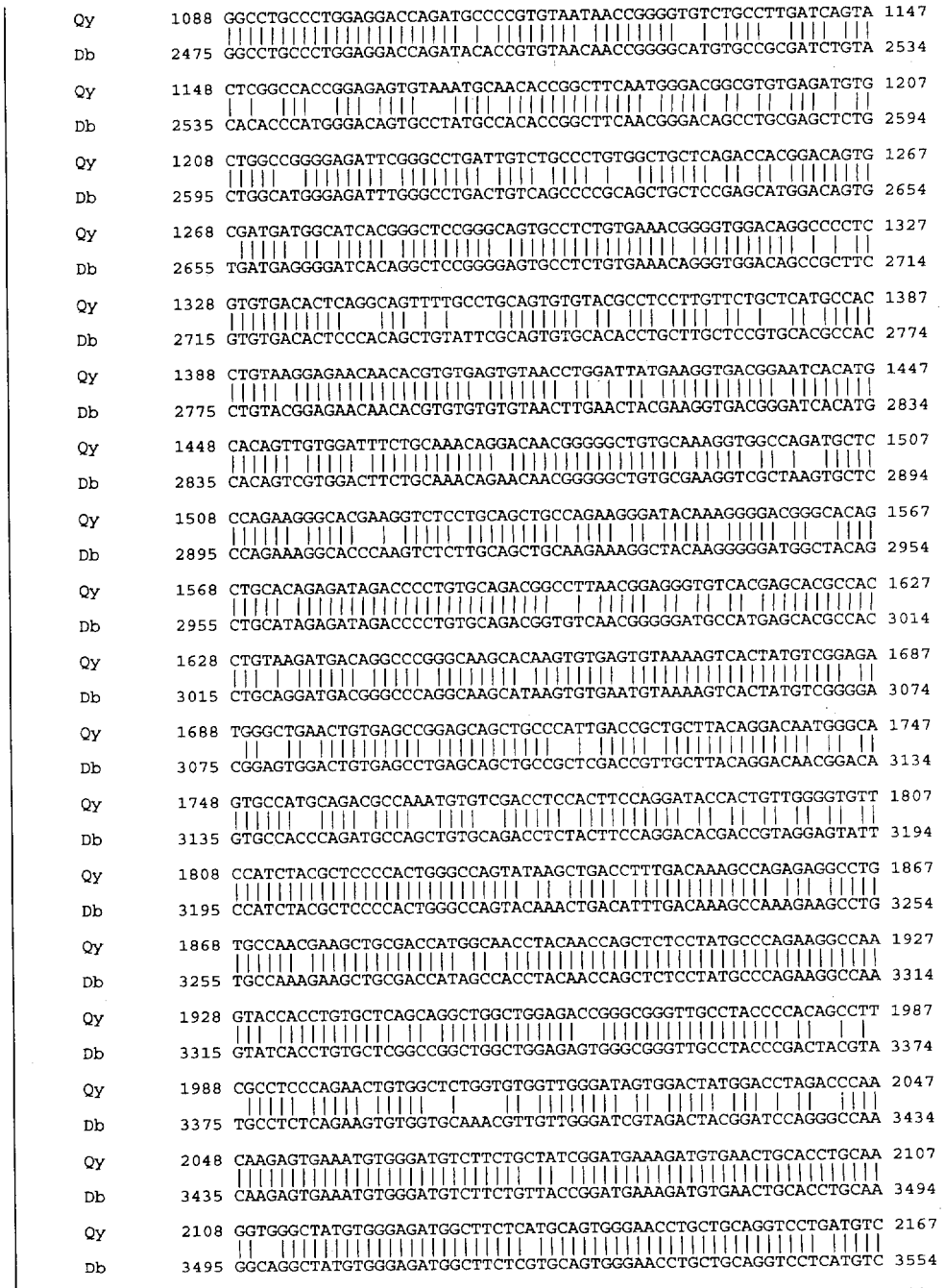
```
CC   acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin
CC   sulphate (CS). HA is an extracellular matrix component of all tissues, in
CC   particular cartilage, skin and vitreous humour. HARE is the endocytic
CC   receptor responsible for removing HA and other glycosaminoglycans from
CC   the circulation. The present sequence is the coding sequence for rat HARE
XX
SQ   Sequence 4706 BP; 1169 A; 1273 C; 1230 G; 1034 T; 0 U; 0 Other;

  Query Match             59.4%;  Score 1936.8;  DB 6;  Length 4706;
  Best Local Similarity   78.3%;  Pred. No. 0;
  Matches 2402;  Conservative    0;  Mismatches  652;  Indels   14;  Gaps    6;

Qy         11 GGAGCTTCCCAAGAACCCGAAAACTTCCCAGTATTTCTTCCAGTTGCAGGAGCATTTCGT 70
Db       1395 GGAGCTTCCCAAGAACCCTTCGACGTCCCAGTACTTCTTCCAGTTGCAGGAGCATGCTGT 1454

Qy         71 GAAAGATCTGGTCGGCCCAGGCCCCTTCACTGTTTTTGCACCTTTATCTGCAGCCTTTGA 130
Db       1455 CCGAGAGCTTGCTGGACCTGGCCCCTTCACCGTGTTCGCGCCTTTGTCTAGCTCCTTCAA 1514

Qy        131 TGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTCGGTA 190
Db       1515 TCATGAGCCCCGGATTAAAGACTGGGATCAGCAGGGCCTCATGTCCCAGGTTCTTCGCTA 1574

Qy        191 CCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTCAAATGCTAC 250
Db       1575 TCACGTGGTGGGCTGCCAGCAGCTGCTGTTGGACAACCTAAAAGTGACCACAAGTGCCAC 1634

Qy        251 TTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTATATAAATAA 310
Db       1635 GACCCTCCAAGGAGAGCCAGTTTCCATCTCTGTCTCTCAGGACACTGTGTTCATAAACAA 1694

Qy        311 TAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCATATCATAGA 370
Db       1695 TGAGGCGAAGGTCCTGTCCAGTGACATCATCAGCACCAATGGCGTCATCCACGTTATAGA 1754

Qy        371 CAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGGAAGAATTCT 430
Db       1755 CAAGTTGCTGTCTCCCAAAAACTTGCTTATCACCCCCAAAGATGCCTTGGGCAGGGTTCT 1814

Qy        431 GCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAACTTAATACA 490
Db       1815 GCAAAATCTTACTACAGTGGCAGCAAACCACGGATATACCAAATTCAGCAAGTTGATACA 1874

Qy        491 GGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCTTCTG 550
Db       1875 GGACTCAGGCTTGCTGTCAGTCATCACTGACTCCATCCACACCCCAGTCACTGTCTTCTG 1934

Qy        551 GCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCAACCA 610
Db       1935 GCCTACGGACAAAGCCCTGGAAGCCTTGCCCCCAGAGCAGCAGGACTTCCTGTTCAATCA 1994

Qy        611 AGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCAAGGT 670
Db       1995 AGACAACAAGGACAAGCTGAAGTCTTACCTGAAGTTCCACGTGATCCGAGACTCCAAGGC 2054

Qy        671 TTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGCTGAG 730
Db       2055 TTTAGCTTCAGACCTCCCCAGGTCTGCTTCCTGGAAGACCCTGCAAGGCTCAGAGCTGAG 2114

Qy        731 TGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCTGCAG 790
Db       2115 TGTGAGGTGTGGAACTGGCAGTGACATCGGTGAGCTCTTTCTAAACGAACAAATGTGCAG 2174

Qy        791 AATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTCTGCT 850
Db       2175 ATTCATACACCGGGGACTCTTGTTTGACGTGGGTGTGGCCTATGGCATTGACTGCCTACT 2234

Qy        851 GATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGGGGGA 910
Db       2235 CATGAATCCTACCCTAGGTGGCCGATGTGACACTTTTACTACCTTCGATATTCCGGGGGA 2294

Qy        911 GTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTGTGAA 970

Db       2295 GTGCGGAAGTTGCATTTTCACTCCCAAATGCCCACTGAAGAGCAAGCCAAAGGGCGTGAA 2354

Qy        971 GCAGAAGTGTCTCTACAACC---TGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGCG 1027
Db       2355 GAAGAAGTGTATCTACAACCCGTTACCTTTCAGGAGGAACGTGGAAGGCTGCCAGAACCT 2414

Qy       1028 GTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTCA 1087
Db       2415 GTGCACCGTGGTGATCCAAACCCCCAGGTGCTGCCATGGTTACTTCATGCCAGACTGTCA 2474

Qy       1088 GGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGTA 1147
Db       2475 GGCCTGCCCTGGAGGACCAGATACACCGTGTAACAACCGGGGCATGTGCCGCGATCTGTA 2534

Qy       1148 CTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGTG 1207
Db       2535 CACACCCATGGGACAGTGCCTATGCCACACCGGCTTCAACGGGACAGCCTGCGAGCTCTG 2594

Qy       1208 CTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGTG 1267
Db       2595 CTGGCATGGGAGATTTGGGCCTGACTGTCAGCCCCGCAGCTGCTCCGAGCATGGACAGTG 2654

Qy       1268 CGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCTC 1327
Db       2655 TGATGAGGGGATCACAGGCTCCGGGGAGTGCCTCTGTGAAACAGGGTGGACAGCCGCTTC 2714

Qy       1328 GTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCAC 1387
Db       2715 GTGTGACACTCCCACAGCTGTATTCGCAGTGTGCACACCTGCTTGCTCCGTGCACGCCAC 2774

Qy       1388 CTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACATG 1447
Db       2775 CTGTACGGAGAACAACACGTGTGTGTGTAACTTGAACTACGAAGGTGACGGGATCACATG 2834

Qy       1448 CACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCTC 1507
Db       2835 CACAGTCGTGGACTTCTGCAAACAGAACAACGGGGGCTGTGCGAAGGTCGCTAAGTGCTC 2894

Qy       1508 CCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACAG 1567
Db       2895 CCAGAAAGGCACCCAAGTCTCTTGCAGCTGCAAGAAAGGCTACAAGGGGGATGGCTACAG 2954

Qy       1568 CTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCAC 1627
Db       2955 CTGCATAGAGATAGACCCCTGTGCAGACGGTGTCAACGGGGGATGCCATGAGCACGCCAC 3014

Qy       1628 CTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAGA 1687
Db       3015 CTGCAGGATGACGGGCCCAGGCAAGCATAAGTGTGAATGTAAAAGTCACTATGTCGGGGA 3074

Qy       1688 TGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGCA 1747
Db       3075 CGGAGTGGACTGTGAGCCTGAGCAGCTGCCGCTCGACCGTTGCTTACAGGACAACGGACA 3134

Qy       1748 GTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGTT 1807
Db       3135 GTGCCACCCAGATGCCAGCTGTGCAGACCTCTACTTCCAGGACACGACCGTAGGAGTATT 3194

Qy       1808 CCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCTG 1867
Db       3195 CCATCTACGCTCCCCACTGGGCCAGTACAAACTGACATTTGACAAAGCCAAAGAAGCCTG 3254

Qy       1868 TGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAA 1927
Db       3255 TGCCAAAGAAGCTGCGACCATAGCCACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAA 3314

Qy       1928 GTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCTT 1987
Db       3315 GTATCACCTGTGCTCGGCCGGCTGGCTGGAGAGTGGGCGGGTTGCCTACCCGACTACGTA 3374

Qy       1988 CGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCAA 2047
```

```
Db       3375 TGCCTCTCAGAAGTGTGGTGCAAACGTTGTTGGGATCGTAGACTACGGATCCAGGGCCAA 3434

Qy       2048 CAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCAA 2107
             |||||||||||||||||||||||||||| || ||||||||||||||||||||||||||||
Db       3435 CAAGAGTGAAATGTGGGATGTCTTCTGTTACCGGATGAAAGATGTGAACTGCACCTGCAA 3494

Qy       2108 GGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGTC 2167
             ||  ||||||||||||||||||||||||| |||||||||||||||||||||||||| |||||
Db       3495 GGCAGGCTATGTGGGAGATGGCTTCTCGTGCAGTGGGAACCTGCTGCAGGTCCTCATGTC 3554

Qy       2168 CTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTCG 2227
             |||||||| |||||||||||||||| || |||||||| | |||||| |||||||| ||
Db       3555 CTTCCCCTCGCTCACAAACTTCCTGACAGAGGTGCTGGCTTTTTCCAAGAGCTCAGCCCG 3614

Qy       2228 AGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCC 2287
             ||| |  |||||| |  |||||||||||||||||||||||| |||||||| ||||||||
Db       3615 AGGACAGGCATTTTTGAAACACCTGACTGACCTGTCCATCCGTGGCACCCTGTTTGTGCC 3674

Qy       2288 ACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACCT 2347
             |||||||||||||||  ||  ||| ||| | ||||||| |||||||| ||||||||||
Db       3675 ACAGAACAGTGGGCTACCGGGAAATAAGAGCCTGTCTGGCCGGGACATTGAGCACCACCT 3734

Qy       2348 CGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGAG 2407
             | | ||||||| | | |  || |||||||||||||||||||| |||   ||| || |
Db       3735 CACTAATGTCAACGTCTCCTTTTACAATGACCTTGTCAATGGTACCTTTCTGAGGACTAT 3794

Qy       2408 GGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCAG 2467
             | |||||||| | |||||||| ||   |||||||||||  ||||| |      ||||||||
Db       3795 GCTGGGAAGCCAACTGCTCATTACCTTCAGCCAGGACCAGCTCCACC---AAGAGACCAG 3851

Qy       2468 GTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCA 2527
             |||||| |||||||||| ||||||||||||||||||||| | ||| ||||||| ||| | ||
Db       3852 GTTTGTGGATGGAAGATCCATTCTGCAGTGGGACATCATCGCCGCCAATGGAATCCTCCA 3911

Qy       2528 TGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCTT 2587
             | | |||||     |||||  | ||| || ||   | | ||     ||||| ||||| |
Db       3912 TATTATTTCTGAACCTTTGAGAGCTCCTCCCACGGCAGCAACGGCTGCCCACTCTGGCCT 3971

Qy       2588 GGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTTA 2647
             |||  |||| || |||| |||| || |||||||  |||||  ||  ||||  |||| |||||
Db       3972 GGGGACAGGTATATTCTGTGCCGTCGTCCTGGTCACTGGTGCGATTGCTCTGGCAGCTTA 4031

Qy       2648 CTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGA 2707
             ||| ||||| ||| |||| | | |||| |   || ||||||| |||||| || ||||||||
Db       4032 CTCTTACTTCCGGCTAAAGCAGCGAACCACTGGTTTCCAGCGTTTTGA-TCAGAAGAGGA 4090

Qy       2708 CATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGA 2767
             |||| ||||    ||| |||||||||||||||  |||||||| | |||||  |||||||
Db       4091 CATTGATGTCTTGGCTTTTGGCAAGCAGCAGCCCAAGAATATCGCAAACCCTCTGTATGA 4150

Qy       2768 GAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACG 2827
             || | || |  | || ||||||||  | ||||  |||||||||||| ||| ||| |||||
Db       4151 GACCTCAGCGCCGGCACCCCCAGAGTCCTCCTGTGACCCCTTCACAGACCCTGGAGAACA 4210

Qy       2828 GCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGC 2887
             | | || |||| ||  |||||  || || |||||    || |  |||  |||| ||||||
Db       4211 GGATCTGGAGGACAGCGACCCTCTGGGGGCACTG---CGGTCCTGACATGAGAAGCCAGC 4267

Qy       2888 CATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGC---ACCAGTTGCCTTTT 2944
             | ||  ||| | |||      |     |  | ||||    ||      |  ||
Db       4268 AAGCAACCACAGTCACGGTTCCACGGTGATTCCCAGCCCCAGCTGTCTCATGGATCAGTT 4327

Qy       2945 AGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGG 3004
                 ||   |   |  ||       |||||||||||| |  ||||| | |||||||
Db       4328 GTTTTAAAGAATGACAACACTCATAAGCCAGCCATACCTCACCCTTCTGGTTAATCTGGG 4387

Qy       3005 GGTTGTTTCTGTGGGTGAGAGATGTGTTG-CTGTGCCCACCCAGTACAGCTTCCTCCTCT 3063
                |  | |  |||   |||||  ||| |      | ||  |  ||||||||
Db       4388 ATTGTCGCCAGGGCTAAGGAGCCATGTTGCCTGGATACCTGGGGGACCTCCACCTCCTCT 4447

Qy       3064 GACCCTTT 3071
              || ||| |
Db       4448 GAGCCTAT 4455
```

RESULT 9
ABX13821
ID ABX13821 standard; cDNA; 4706 BP.
XX
AC ABX13821;
XX
DT 19-FEB-2003 (first entry)
XX
DE cDNA encoding rat 175kDa Hyaluronan receptor for endocytosis (HARE).
XX
KW Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
KW tumour; gene therapy; rat; gene; ss.
XX
OS Rattus sp.
XX
FH Key   Location/Qualifiers
FT CDS   1. .4296
FT   /*tag= a
FT   /product= "HARE"
FT   /note= "Hyaluronan receptor for endocytosis"
FT   /partial
FT   /note= "No start codon given"
XX
PN WO200286093-A2.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013209.
XX
PR 25-APR-2001; 2001US-00842930.
PR 25-APR-2001; 2001US-0286468P.
XX
PA (WEIG/) WEIGEL P H.
PA (WEIG/) WEIGEL J A.
XX
PI Weigel PH, Weigel JA;
XX
DR WPI; 2003-093126/08.
DR P-PSDB; ABG72498.
XX
PT Targeting compounds e.g. chemotherapeutic agent to cell of subject
PT expressing functional active hyaluronan receptor for endocytosis of HARE,
PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
PT epitope.
XX
PS Example; Fig 2; 167pp; English.
XX
CC The invention describes a method of targeting a compound to a cell or
CC tissue of an individual expressing a functionally active hyaluronan (HA)
CC receptor for endocytosis (HARE) or a cell that does not express
CC functionally active HARE. The method involves using HA molecule, a
CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
CC a monoclonal antibody raised against a HA-binding domain of HARE. The
CC method is useful for targeting a compound, preferably a chemotherapeutic
CC agent or a radioisotope to cell of an individual, especially a human,
CC expressing HARE on its surface (e.g. gene therapy). Also described is a
CC method useful for preventing interaction between a cell having at least
CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
CC its surface. This second method is useful for preventing metastasis by
CC preventing interaction between tumour cells having HA, CD or CDS coat and
CC non-tumour cells expressing HARE on its surface. The invention also
CC describes a method useful for detecting the presence of HA, CD and CDS in
CC a biological fluid. This sequence encodes the rat 175kDa Hyaluronan
CC receptor for endocytosis (HARE)
XX
SQ Sequence 4706 BP; 1168 A; 1273 C; 1230 G; 1035 T; 0 U; 0 Other;

Query Match 59.4%; Score 1936.8; DB 7; Length 4706;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 652; Indels 14; Gaps 6;





```
Db       3555 CTTCCCCTCGCTCACAAACTTCCTGACAGAGGTGCTGGCTTTTTCCAAGAGCTCAGCCCG 3614

Qy       2228 AGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCC 2287

Db       3615 AGGACAGGCATTTTTGAAACACCTGACTGACCTGTCCATCCGTGGCACCCTGTTTGTGCC 3674

Qy       2288 ACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACCT 2347

Db       3675 ACAGAACAGTGGGCTACCGGGAAATAAGAGCCTGTCTGGCCGGGACATTGAGCACCACCT 3734

Qy       2348 CGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGAG 2407

Db       3735 CACTAATGTCAACGTCTCCTTTTACAATGACCTTGTCAATGGTACCTTTCTGAGGACTAT 3794

Qy       2408 GGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCGACGGAGACCAG 2467

Db       3795 GCTGGGAAGCCAACTGCTCATTACCTTCAGCCAGGACCAGCTCCACC---AAGAGACCAG 3851

Qy       2468 GTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCA 2527

Db       3852 GTTTGTGGATGGAAGATCCATTCTGCAGTGGGACATCATCGCCGCCAATGGAATCCTCCA 3911

Qy       2528 TGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCTT 2587

Db       3912 TATTATTTCTGAACCTTTGAGAGCTCCTCCCACGGCAGCAACGGCTGCCCACTCTGGCCT 3971

Qy       2588 GGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTTA 2647

Db       3972 GGGGACAGGTATATTCTGTGCCGTCGTCCTGGTCACTGGTGCGATTGCTCTGGCAGCTTA 4031

Qy       2648 CTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGA 2707

Db       4032 CTCTTACTTCCGGCTAAAGCAGCGAACCACTGGTTTCCAGCGTTTTGA-TCAGAAGAGGA 4090

Qy       2708 CATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGA 2767

Db       4091 CATTGATGTCTTGGCTTTTGGCAAGCAGCAGCCCAAGAATATCGCAAACCCTCTGTATGA 4150

Qy       2768 GAGCACAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACG 2827

Db       4151 GACCTCAGCGCCGGCACCCCCAGAGTCCTCCTGTGACCCCTTCACAGACCCTGGAGAACA 4210

Qy       2828 GCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGC 2887

Db       4211 GGATCTGGAGGACAGCGACCCTCTGGGGGCACTG---CGGTCCTGACATGAGAAGCCAGC 4267

Qy       2888 CATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGC---ACCAGTTGCCTTTT 2944

Db       4268 AAGCAACCACAGTCACGGTTCCACGGTGATTCCCAGCCCCAGCTGTCTCATGGATCAGTT 4327

Qy       2945 AGGAACGTAAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGG 3004

Db       4328 GTTTTAAAGAATGACAACACTCATAAGCCAGCCATACCTCACCCTTCTGGTTAATCTGGG 4387

Qy       3005 GGTTGTTTCTGTGGGTGAGAGATGTGTTG-CTGTGCCCACCCAGTACAGCTTCCTCCTCT 3063

Db       4388 ATTGTCGCCAGGGCTAAGGAGCCATGTTGCCTGGATACCTGGGGGACCTCCACCTCCTCT 4447

Qy       3064 GACCCTTT 3071

Db       4448 GAGCCTAT 4455


RESULT 10
AAF87119
ID   AAF87119 standard; DNA; 2483 BP.
XX
AC   AAF87119;
XX
DT   26-MAR-2002  (first entry)
XX
DE   NOV8 coding sequence.
XX
KW   NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW   cardiovascular; casein kinase II phosphorylation site; contraception;
KW   serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW   epidermal growth factor; cell development; apoptosis; cell adhesion;
KW   growth migration; cell structure; motility; cancer; immune disorder;
KW   inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
KW   cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW   therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW   NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             387. .2396
FT                   /*tag=  a
FT                   /product= "NOV8"
XX
PN   WO200136638-A2.
XX
PD   25-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US031543.
XX
PR   19-NOV-1999;   99US-0166336P.
PR   29-NOV-1999;   99US-0167785P.
PR   08-MAR-2000; 2000US-0187844P.
PR   16-NOV-2000; 2000US-00715417.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Lichenstein H,  Vernet C,  Fernandes E;
XX
DR   WPI; 2001-648134/74.
DR   P-PSDB; AAB83364.
XX
PT   Novel human polypeptides and the nucleic acids that encode them useful
PT   for preventing, diagnosing and treating e.g. cancer, inflammation and
PT   immune disorders.
XX
PS   Claim 9; Page 29-30; 141pp; English.
XX
CC   This sequence encodes the NOV8 protein. The invention relates to the NOV1
CC   -NOV16 proteins, and their coding sequences. The proteins have Cytostatic
CC   ; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
CC   activities. The sequences may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate NOVX expression. They
CC   may be used to treat disorders associated with decreased expression by
CC   rectifying mutations or deletions in a patient's genome that affect the
CC   activity of protein by expressing inactive proteins or to supplement the
CC   patients own production of protein. They are used to produce NOVX
CC   proteins, by inserting the nucleic acid into a cell and culturing it to
CC   express the protein. The DNA may be used as DNA probes in assays to
CC   detect and quantitate the presence of similar DNAs in samples, and which
CC   patients may need restorative therapy. The NOVX protein may also be used
CC   as antigens in the production of antibodies (Abs) against NOVX and in
CC   assays to identify modulators of NOVX expression and activity. The anti-
CC   NOVX Abs and antagonist are used to down regulate expression and
CC   activity. The anti-NOVX Abs are used for detecting the presence of NOVX
CC   in samples. Disorders that may be prevented, diagnosed and/or treated
CC   vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
CC   NOV13-16 have casein kinase II phosphorylation sites characteristic of
CC   serine/threonine kinases, and are used to treat kinase-related disorders
CC   (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
CC   NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
CC   -like super family and are involved in, e.g. regulation of cell
CC   development, apoptosis, cell adhesion, growth migration, cell structure
CC   and motility and protein management, and are used to treat cancers,
CC   inflammatory disorders, immune disorders and cellular adhesion disorders.
CC   NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
CC   treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
CC   syndrome and marfan syndrome
XX
```

```
SQ   Sequence 2483 BP; 619 A; 648 C; 639 G; 577 T; 0 U; 0 Other;

  Query Match             54.6%;  Score 1778.6;  DB 4;  Length 2483;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1790;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy        122 AGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGT 181
              | ||| ||||          | ||||||||||||||||||||||||||||||||||||||||||
Db        569 ACCCTCTGATCTTCTGTTACAGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGT 628

Qy        182 TCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTC 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTC 688

Qy        242 AAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 AAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTA 748

Qy        302 TATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCA 361
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 TATAAACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCA 808

Qy        362 TATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 TATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGG 868

Qy        422 AAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAA 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 AAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAA 928

Qy        482 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCAC 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCAC 988

Qy        542 TCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCT 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 TCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCT 1048

Qy        602 GTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 GTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 1108

Qy        662 TGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTC 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTC 1168

Qy        722 AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCA 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCA 1228

Qy        782 AACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGA 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 AACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGA 1288

Qy        842 CTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGC 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 CTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGC 1348

Qy        902 CTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAA 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1349 CTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAA 1408

Qy        962 GGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCG 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 GGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCG 1468

Qy       1022 GGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGA 1081
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1469 GGAGCGATGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGA 1528

Qy       1082 CTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGA 1141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1529 CTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGA 1588

Qy       1142 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA 1201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1589 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA 1648

Qy       1202 GATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGG 1261
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1649 GATGTGCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGG 1708

Qy       1262 ACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGG 1321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 ACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGG 1768

Qy       1322 CCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCA 1381
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       1769 CCCCTCGTGTGACACTCAGGCAGTTTTGTCTGCAGTGTGTACGCCTCCTTGTTCTGCTCA 1828

Qy       1382 TGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAAT 1441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1829 TGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAAT 1888

Qy       1442 CACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAG 1501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1889 CACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAG 1948

Qy       1502 ATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGG 1561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1949 ATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGG 2008

Qy       1562 GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCA 1621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2009 GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCA 2068

Qy       1622 CGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGT 1681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2069 CGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGT 2128

Qy       1682 CGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAA 1741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2129 CGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAA 2188

Qy       1742 TGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGG 1801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2189 TGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGG 2248

Qy       1802 GGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2249 GGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA 2308

Qy       1862 GGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAA 1921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2309 GGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAA 2368

Qy       1922 GGCCAAGTA 1930
              | ||  |||
Db       2369 GACCTGGTA 2377


RESULT 11
AAF87117
ID   AAF87117 standard; DNA; 2512 BP.
XX
AC   AAF87117;
XX
DT   26-MAR-2002  (first entry)
XX
DE   NOV6 coding sequence.
XX
KW   NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW   cardiovascular; casein kinase II phosphorylation site; contraception;
KW   serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW   epidermal growth factor; cell development; apoptosis; cell adhesion;
KW   growth migration; cell structure; motility; cancer; immune disorder;
```

```
KW   inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
KW   cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW   therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW   NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             416. .2425
FT                   /*tag= a
FT                   /product= "NOV6"
XX
PN   WO200136638-A2.
XX
PD   25-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US031543.
XX
PR   19-NOV-1999;   99US-0166336P.
PR   29-NOV-1999;   99US-0167785P.
PR   08-MAR-2000; 2000US-0187844P.
PR   16-NOV-2000; 2000US-00715417.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Lichenstein H,  Vernet C,  Fernandes E;
XX
DR   WPI; 2001-648134/74.
DR   P-PSDB; AAB83362.
XX
PT   Novel human polypeptides and the nucleic acids that encode them useful
PT   for preventing, diagnosing and treating e.g. cancer, inflammation and
PT   immune disorders.
XX
PS   Claim 9; Page 24-25; 141pp; English.
XX
CC   This sequence encodes the NOV6 protein. The invention relates to the NOV1
CC   -NOV16 proteins, and their coding sequences. The proteins have Cytostatic
CC   ; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
CC   activities. The sequences may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate NOVX expression. They
CC   may be used to treat disorders associated with decreased expression by
CC   rectifying mutations or deletions in a patient's genome that affect the
CC   activity of protein by expressing inactive proteins or to supplement the
CC   patients own production of protein. They are used to produce NOVX
CC   proteins, by inserting the nucleic acid into a cell and culturing it to
CC   express the protein. The DNA may be used as DNA probes in assays to
CC   detect and quantitate the presence of similar DNAs in samples, and which
CC   patients may need restorative therapy. The NOVX protein may also be used
CC   as antigens in the production of antibodies (Abs) against NOVX and in
CC   assays to identify modulators of NOVX expression and activity. The anti-
CC   NOVX Abs and antagonist are used to down regulate expression and
CC   activity. The anti-NOVX Abs are used for detecting the presence of NOVX
CC   in samples. Disorders that may be prevented, diagnosed and/or treated
CC   vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
CC   NOV13-16 have casein kinase II phosphorylation sites characteristic of
CC   serine/threonine kinases, and are used to treat kinase-related disorders
CC   (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
CC   NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
CC   -like super family and are involved in, e.g. regulation of cell
CC   development, apoptosis, cell adhesion, growth migration, cell structure
CC   and motility and protein management, and are used to treat cancers,
CC   inflammatory disorders, immune disorders and cellular adhesion disorders.
CC   NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
CC   treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
CC   syndrome and marfan syndrome
XX
SQ   Sequence 2512 BP; 623 A; 657 C; 651 G; 581 T; 0 U; 0 Other;

  Query Match           54.5%;  Score 1777;  DB 4;  Length 2512;
  Best Local Similarity  98.9%;  Pred. No. 0;
  Matches 1789;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       122 AGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGT 181
Db       598 ACCCTCTGATCTTCTGTTACAGGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGT 657

Qy       182 TCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTC 241
Db       658 TCTTCGGTACCATGTGGTCGCCTGCCACCAGCTGCTTCTGGAAAACCTGAAATTGATCTC 717

Qy       242 AAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTA 301
Db       718 AAATGCTACTTCCCTCCAAGGAGAGCCAATAGTCATCTCCGTCTCTCAGAGCACGGTGTA 777

Qy       302 TATAAATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCA 361
Db       778 TATAAACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTGTTCA 837

Qy       362 TATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGG 421
Db       838 TATCATAGACAAATTGCTATCTCCCAAAAATTTGCTTATCACTCCCAAAGACAACTCTGG 897

Qy       422 AAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAA 481
Db       898 AAGAATTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAATTTAGCAA 957

Qy       482 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCAC 541
Db       958 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCAC 1017

Qy       542 TCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCT 601
Db      1018 TCTCTTCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCT 1077

Qy       602 GTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 661
Db      1078 GTTCAACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 1137

Qy       662 TGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTC 721
Db      1138 TGCCAAGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTC 1197

Qy       722 AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCA 781
Db      1198 AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCA 1257

Qy       782 AACCTGCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGA 841
Db      1258 AACCTGCAGAATTGTGCAGCGCGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGA 1317

Qy       842 CTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGC 901
Db      1318 CTGTCTGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGC 1377

Qy       902 CTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAA 961
Db      1378 CTCGGGGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAA 1437

Qy       962 GGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCG 1021
Db      1438 GGGTGTGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCG 1497

Qy      1022 GGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGA 1081
Db      1498 GGAGCGGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGA 1557

Qy      1082 CTGTCAGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGA 1141
Db      1558 CTGTCAGGCCTGCCCTGGAGGACCAGTTGCCCCGTGTAATAACCGGGGTGTCTGCCTTGA 1617

Qy      1142 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA 1201
Db      1618 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA 1677

Qy      1202 GATGTGCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGG 1261
```

```
Db       1678 GATGTGCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGG 1737

Qy       1262 ACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGG 1321

Db       1738 ACAGTGCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGG 1797

Qy       1322 CCCCTCGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCA 1381

Db       1798 CCCCTCGTGTGACACTCAGGCAGTTTTGTCTGCAGTGTGTACGCCTCCTTGTTCTGCTCA 1857

Qy       1382 TGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAAT 1441

Db       1858 TGCCACCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAAT 1917

Qy       1442 CACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAG 1501

Db       1918 CACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAG 1977

Qy       1502 ATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGG 1561

Db       1978 ATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGG 2037

Qy       1562 GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCA 1621

Db       2038 GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCA 2097

Qy       1622 CGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGT 1681

Db       2098 CGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGT 2157

Qy       1682 CGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAA 1741

Db       2158 CGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAA 2217

Qy       1742 TGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGG 1801

Db       2218 TGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGG 2277

Qy       1802 GGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA 1861

Db       2278 GGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA 2337

Qy       1862 GGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAA 1921

Db       2338 GGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAA 2397

Qy       1922 GGCCAAGTA 1930

Db       2398 GACCTGGTA 2406


RESULT 12
AAF87113
ID   AAF87113 standard; DNA; 2011 BP.
XX
AC   AAF87113;
XX
DT   26-MAR-2002  (first entry)
XX
DE   NOV2 coding sequence.
XX
KW   NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW   cardiovascular; casein kinase II phosphorylation site; contraception;
KW   serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW   epidermal growth factor; cell development; apoptosis; cell adhesion;
KW   growth migration; cell structure; motility; cancer; immune disorder;
KW   inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
KW   cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW   therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW   NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             877. .1824
FT                   /*tag= a
FT                   /product= "NOV2"
XX
PN   WO200136638-A2.
XX
PD   25-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US031543.
XX
PR   19-NOV-1999;   99US-0166336P.
PR   29-NOV-1999;   99US-0167785P.
PR   08-MAR-2000; 2000US-0187844P.
PR   16-NOV-2000; 2000US-00715417.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Lichenstein H,  Vernet C,  Fernandes E;
XX
DR   WPI; 2001-648134/74.
DR   P-PSDB; AAB83358.
XX
PT   Novel human polypeptides and the nucleic acids that encode them useful
PT   for preventing, diagnosing and treating e.g. cancer, inflammation and
PT   immune disorders.
XX
PS   Claim 9; Page 10-13; 141pp; English.
XX
CC   This sequence encodes the NOV2 protein. The invention relates to the NOV1
CC   -NOV16 proteins, and their coding sequences. The proteins have Cytostatic
CC   ; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
CC   activities. The sequences may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate NOVX expression. They
CC   may be used to treat disorders associated with decreased expression by
CC   rectifying mutations or deletions in a patient's genome that affect the
CC   activity of protein by expressing inactive proteins or to supplement the
CC   patients own production of protein. They are used to produce NOVX
CC   proteins, by inserting the nucleic acid into a cell and culturing it to
CC   express the protein. The DNA may be used as DNA probes in assays to
CC   detect and quantitate the presence of similar DNAs in samples, and which
CC   patients may need restorative therapy. The NOVX protein may also be used
CC   as antigens in the production of antibodies (Abs) against NOVX and in
CC   assays to identify modulators of NOVX expression and activity. The anti-
CC   NOVX Abs and antagonist are used to down regulate expression and
CC   activity. The anti-NOVX Abs are used for detecting the presence of NOVX
CC   in samples. Disorders that may be prevented, diagnosed and/or treated
CC   vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
CC   NOV13-16 have casein kinase II phosphorylation sites characteristic of
CC   serine/threonine kinases, and are used to treat kinase-related disorders
CC   (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
CC   NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
CC   -like super family and are involved in, e.g. regulation of cell
CC   development, apoptosis, cell adhesion, growth migration, cell structure
CC   and motility and protein management, and are used to treat cancers,
CC   inflammatory disorders, immune disorders and cellular adhesion disorders.
CC   NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
CC   treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
CC   syndrome and marfan syndrome
XX
SQ   Sequence 2011 BP; 458 A; 551 C; 583 G; 419 T; 0 U; 0 Other;

  Query Match             52.8%;  Score 1721.2;  DB 4;  Length 2011;
  Best Local Similarity   94.1%;  Pred. No. 0;
  Matches 1843;  Conservative    0;  Mismatches    8;  Indels  107;  Gaps    1;

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546

Db        161 TTCAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 220

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
```

```
Db        221 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 280
Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
Db        281 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 340
Qy        667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
Db        341 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 400
Qy        727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
Db        401 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 460
Qy        787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
Db        461 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 520
Qy        847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
Db        521 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 580
Qy        907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
Db        581 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 640
Qy        967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
Db        641 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 700
Qy       1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
Db        701 GATGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 760
Qy       1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
Db        761 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 820
Qy       1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
Db        821 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 880
Qy       1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
Db        881 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 940
Qy       1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
Db        941 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1000
Qy       1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
Db       1001 CGTGTGACACTCAGGCAGTTTTGTCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1060
Qy       1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
Db       1061 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1120
Qy       1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
Db       1121 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1180
Qy       1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
Db       1181 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1240
Qy       1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
Db       1241 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1300
Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
Db       1301 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1360
```

```
Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
Db       1361 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1420
Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
Db       1421 AGTGCCATGCAGACGCCAAATGTGCCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1480
Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
Db       1481 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1540
Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1926
Db       1541 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCA 1600
Qy       1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1986
Db       1601 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCT 1660
Qy       1987 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCA 2046
Db       1661 TCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCCAGACCCA 1720
Qy       2047 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCA 2106
Db       1721 ACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAA------------------- 1761
Qy       2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGT 2166
Db       1762 ------------------------------------------------------------ 1761
Qy       2167 CCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 2226
Db       1762 ----------------------------GGAAGTGCTGGCCTATTCCAACAGCTCAGCTC 1793
Qy       2227 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 2286
Db       1794 GAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGC 1853
Qy       2287 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACC 2346
Db       1854 CACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAACACCACC 1913
Qy       2347 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 2406
Db       1914 TCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACCCTGCAAACGA 1973
Qy       2407 GGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGAC 2444
Db       1974 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGAC 2011
```

RESULT 13
AAF87114
ID AAF87114 standard; DNA; 1804 BP.
XX
AC AAF87114;
XX
DT 26-MAR-2002 (first entry)
XX
DE NOV3 coding sequence.
XX
KW NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW cardiovascular; casein kinase II phosphorylation site; contraception;
KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW epidermal growth factor; cell development; apoptosis; cell adhesion;
KW growth migration; cell structure; motility; cancer; immune disorder;
KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.

```
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             877. .1611
FT                   /*tag=  a
FT                   /product= "NOV3"
XX
PN   WO200136638-A2.
XX
PD   25-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US031543.
XX
PR   19-NOV-1999;   99US-0166336P.
PR   29-NOV-1999;   99US-0167785P.
PR   08-MAR-2000; 2000US-0187844P.
PR   16-NOV-2000; 2000US-00715417.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Lichenstein H,  Vernet C,  Fernandes E;
XX
DR   WPI; 2001-648134/74.
DR   P-PSDB; AAB93359.
XX
PT   Novel human polypeptides and the nucleic acids that encode them useful
PT   for preventing, diagnosing and treating e.g. cancer, inflammation and
PT   immune disorders.
XX
PS   Claim 9; Page 14-17; 141pp; English.
XX
CC   This sequence encodes the NOV3 protein. The invention relates to the NOV1
CC   -NOV16 proteins, and their coding sequences. The proteins have Cytostatic
CC   ; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
CC   activities. The sequences may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate NOVX expression. They
CC   may be used to treat disorders associated with decreased expression by
CC   rectifying mutations or deletions in a patient's genome that affect the
CC   activity of protein by expressing inactive proteins or to supplement the
CC   patients own production of protein. They are used to produce NOVX
CC   proteins, by inserting the nucleic acid into a cell and culturing it to
CC   express the protein. The DNA may be used as DNA probes in assays to
CC   detect and quantitate the presence of similar DNAs in samples, and which
CC   patients may need restorative therapy. The NOVX protein may also be used
CC   as antigens in the production of antibodies (Abs) against NOVX and in
CC   assays to identify modulators of NOVX expression and activity. The anti-
CC   NOVX Abs and antagonist are used to down regulate expression and
CC   activity. The anti-NOVX Abs are used for detecting the presence of NOVX
CC   in samples. Disorders that may be prevented, diagnosed and/or treated
CC   vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
CC   NOV13-16 have casein kinase II phosphorylation sites characteristic of
CC   serine/threonine kinases, and are used to treat kinase-related disorders
CC   (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
CC   NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
CC   -like super family and are involved in, e.g. regulation of cell
CC   development, apoptosis, cell adhesion, growth migration, cell structure
CC   and motility and protein management, and are used to treat cancers,
CC   inflammatory disorders, immune disorders and cellular adhesion disorders.
CC   NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
CC   treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
CC   syndrome and marfan syndrome
XX
SQ   Sequence 1804 BP; 426 A; 487 C; 521 G; 370 T; 0 U; 0 Other;

  Query Match             43.9%;  Score 1429.6;  DB 4;  Length 1804;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1432;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        487 TACAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 546
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TTCAGGACTCAGGTTTGCTGAGTGTCATCACCGATCCCATCCACACCCCAGTCACTCTCT 220

Qy        547 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 TCTGGCCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCCTGTTCA 280

Qy        607 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 ACCAAGACAACAAGGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATACGAGATGCCA 340

Qy        667 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AGGTTTTAGCTGTGGATCTTCCCACATCCACTGCCTGGAAGACCCTGCAAGGTTCAGAGC 400

Qy        727 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 TGAGTGTGAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 460

Qy        787 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GCAGAATTGTGCAGCGGGAGCTCTTGTTTGACCTGGGTGTGGCCTACGGCATTGACTGTC 520

Qy        847 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 TGCTGATTGATCCCACCCTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATGCCTCGG 580

Qy        907 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 GGGAGTGTGGGAGCTGTGTCAATACTCCCAGCTGCCCAAGGTGGAGTAAACCAAAGGGTG 640

Qy        967 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC 700

Qy       1027 GGTGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 1086
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 GATGCAGCCTGGTGATACAGATCCCCAGGTGCTGCAAGGGCTACTTCGGGCGAGACTGTC 760

Qy       1087 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 AGGCCTGCCCTGGAGGACCAGATGCCCCGTGTAATAACCGGGGTGTCTGCCTTGATCAGT 820

Qy       1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 880

Qy       1207 GCTGGCCGGGGAGATTCGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 1266
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        881 GCTGGCCGGGGAGATTTGGGCCTGATTGTCTGCCCTGTGGCTGCTCAGACCACGGACAGT 940

Qy       1267 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 GCGATGATGGCATCACGGGCTCCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCCT 1000

Qy       1327 CGTGTGACACTCAGGCAGTTTTGCCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1386
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1001 CGTGTGACACTCAGGCAGTTTTGTCTGCAGTGTGTACGCCTCCTTGTTCTGCTCATGCCA 1060

Qy       1387 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1061 CCTGTAAGGAGAACAACACGTGTGAGTGTAACCTGGATTATGAAGGTGACGGAATCACAT 1120

Qy       1447 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1121 GCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTGGCCAGATGCT 1180

Qy       1507 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1181 CCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGGGACGGGCACA 1240

Qy       1567 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1241 GCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCACGAGCACGCCA 1300

Qy       1627 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1686
```

```
Db       1301 CCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCACTATGTCGGAG 1360

Qy       1687 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1746
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1361 ATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAGGACAATGGGC 1420

Qy       1747 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1806
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1421 AGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACTGTTGGGGTGT 1480

Qy       1807 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1481 TCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGGCCT 1540

Qy       1867 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAG 1922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1541 GTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAG 1596
```

RESULT 14
AAC76373
ID AAC76373 standard; cDNA; 1377 BP.
XX
AC AAC76373;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1928 polynucleotide sequence SEQ ID NO:3855.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US008621.
XX
PR 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR P-PSDB; AAB42164.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 5; Page 3006-3007; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 1377 BP; 303 A; 398 C; 342 G; 330 T; 0 U; 4 Other;

Query Match 41.3%; Score 1347.4; DB 3; Length 1377;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1371; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

```
Qy       1876 AAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAAGTACCACC 1935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCCCAGAAGGCCAAGTACCACC 61

Qy       1936 TGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCTTCGCCTCCC 1995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTACCCCACAGCCTTCGCCTCCC 121

Qy       1996 AGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCAACAAGAGTG 2055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGACCTAGACCCAACAAGAGTG 181

Qy       2056 AAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCAAGGTGGGCT 2115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 AAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACTGCACCTGCAAGGTGGGCT 241

Qy       2116 ATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGTCCTTCCCCT 2175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAGGTCCTGATGTCCTTCCCCT 301

Qy       2176 CACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTCGAGGCCGTG 2235
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAACAGCTCAGCTCGAGGCCGTG 361

Qy       2236 CATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCCACAGAACA 2295
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACCCTCTTTGTGCCACAGAACA 421

Qy       2296 GTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACCTCGCCAATG 2355
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATCGAGCACCACCTCGCCAATG 481

Qy       2356 TCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACC-ACCCTGCAAACGAGGGTGGGA 2414
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        482 TCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCNACCCTGCAAACGAGGGTGGGA 541

Qy       2415 AGCAAGCTGCTCATCACTGCCAGCCAGGACCCACT--CCAACCGACGGAGACCAGGTTTG 2472
             |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        542 AGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTNNCCAACCGACGGAGACCAGGTTTG 601

Qy       2473 TTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCATGTCA 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 TTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAATGGGATCATTCATGTCA 661

Qy       2533 TTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCTTGGGAG 2592
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACCCACACTGGCTTGGGAG 721
```

```
Qy       2593 CAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTTACTCCT 2652
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 CAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCCTTGGCTGCTTACTCCT 781

Qy       2653 ACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGACATTA 2712
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 ACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAGGACATTA 841

Qy       2713 ATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGAGAGCA 2772
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ATGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAACCCCTTGTATGAGAGCA 901

Qy       2773 CAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGC 2832
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 CAACCTCAGCTCCCCCAGAACCTTCCTACGACCCCTTCACGGACTCTGAAGAACGGCAGC 961

Qy       2833 TTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCA 2892
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 TTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGGACGGGAGATGCCAGCCATCA 1021

Qy       2893 CTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGT 2952
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAGTTGCCTTTTAGGAACGT 1081

Qy       2953 AAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTT 3012
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 AAAGTCCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGGGGTTGTTT 1141

Qy       3013 CTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTG 3072
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 CTGTGGGTGAGAGATGTGTTGCTGTGCCCACCCAGTACAGCTTCCTCCTCTGACCCTTTG 1201

Qy       3073 GCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGT 3132
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 GCTCTTCTTCCTTTGTACTCTTCAGCTGGCACCTGCTCCATTCTGCCCTACATGATGGGT 1261

Qy       3133 AACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCNTCTTTGTATCCCAGCCCCTAG 3192
             ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       1262 AACTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGTCTTTGTATCCCAGCCCCTAG 1321

Qy       3193 CCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAAAAAAA 3248
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1322 CCCAGTGCCTGACACAGGAACTGTGCACAATAAAGGTTTATGGAACAGAAACAAAA 1377
```

RESULT 15
AAA57365
ID AAA57365 standard; DNA; 1259 BP.
XX
AC AAA57365;
XX
DT 03-OCT-2000 (first entry)
XX
DE DNA encoding a human hyaluronan-binding protein, designated BM-HABP.
XX
KW Hyaluronan-binding protein WF-HABP; OE-HABP; BM-HABP;
KW proliferative condition; metastasis; inflammation; ischemia;
KW host defence dysfunction; immune surveillance dysfunction; arthritis;
KW multiple sclerosis; autoimmunity; immune dysfunction; allergy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 199. .1257
FT /*tag= a
FT /product= "hyaluronan-binding protein"
FT /transl_except= (pos: 478. .480, aa: Xaa)
FT /transl_except= (pos: 667. .669, aa: Xaa)
FT /transl_except= (pos: 1105. .1107, aa: Xaa)
FT /transl_except= (pos: 1138. .1140, aa: Xaa)
FT /transl_except= (pos: 1156. .1158, aa: Xaa)
FT /transl_except= (pos: 1168. .1170, aa: Xaa)
FT /transl_except= (pos: 1171. .1173, aa: Xaa)
FT /transl_except= (pos: 1249. .1251, aa: Xaa)
FT /note= "Xaa is an unspecified amino acid; no termination
FT codon given"
XX
PN WO200039166-A1.
XX
PD 06-JUL-2000.
XX
PF 20-DEC-1999; 99WO-US030462.
XX
PR 23-DEC-1998; 98US-0113871P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (AMNA-) AMERICAN NAT RED CROSS.
XX
PI Hastings GA, Liau G, Tsifrina E;
XX
DR WPI; 2000-452376/39.
DR P-PSDB; AAY93913.
XX
PT New hyaluron-binding proteins, known as full-length WF-HABP, WF-HABP, OE-
PT HABP and BM-HABP, useful for treating proliferative conditions,
PT metastasis, inflammation, ischemia, arthritis and multiple sclerosis.
XX
PS Claim 2; Fig 4A-B; 457pp; English.
XX
CC The present sequence encodes a hyaluronan-binding protein. The
CC specification describes four hyaluronan-binding protein, known as WF-
CC HABP, WF-HABP, OE-HABP, and BM-HABP. The polypeptides are useful for
CC treating diseases such as proliferative conditions, metastasis,
CC inflammation, ischemia, host defence dysfunction, immune surveillance
CC dysfunction, arthritis, multiple sclerosis, autoimmunity, immune
CC dysfunction and allergy
XX
SQ Sequence 1259 BP; 295 A; 348 C; 346 G; 258 T; 0 U; 12 Other;

Query Match 37.7%; Score 1227.4; DB 3; Length 1259;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1234; Conservative 2; Mismatches 23; Indels 0; Gaps 0;

```
Qy       1437 GGAATCACATGCACAGTTGTGGATTTCTGCAAACAGGACAACGGGGGCTGTGCAAAGGTG 1496
             ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db          1 GGAATCACATGCACAGTTGTGGATTTYTGCAAACAGGACAACGGGGGCTGTGCAAAGGTG 60

Qy       1497 GCCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGG 120

Qy       1557 GACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCAC 1616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTCAC 180

Qy       1617 GAGCACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCAC 1676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCAC 240

Qy       1677 TATGTCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAG 1736
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TATGTCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAG 300

Qy       1737 GACAATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACT 1796
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACAATGGGCAGTGCCATGCAGACGCCAAATGTGTCGACCTCCACTTCCAGGATACCACT 360

Qy       1797 GTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCC 1856
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTTGGGGTGTTCCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCC 420

Qy       1857 AGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATGCC 1916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        421 AGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATNNC 480

Qy       1917 CAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTAC 1976
```

```
Db        481 CAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTAC 540

Qy       1977 CCCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGA 2036

Db        541 CCCACAGCCTTCGCCTCCCAGAACTGTGGCTCTGGTGTGGTTGGGATAGTGGACTATGGA 600

Qy       2037 CCTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAAC 2096

Db        601 CCTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAAC 660

Qy       2097 TGCACCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACCTGCTGCAG 2156

Db        661 TGCACCTNCAAGGTGGGCTATGTGGGAGATGGCTTCTCATACAGTGGGAACCTGCTGCAG 720

Qy       2157 GTCCTGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAAC 2216

Db        721 GTCCTGATGTCCTTCCCCTCACTCACAAACTTCCTGACGGAAGTGCTGGCCTATTCCAAC 780

Qy       2217 AGCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACC 2276

Db        781 AGCTCAGCTCGAGGCCGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGCACC 840

Qy       2277 CTCTTTGTGCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATC 2336

Db        841 CTCTTTGTNCCACAGAACAGTGGGCTGGGGGAGAATGAGACCTTGTCTGGGCGGGACATC 900

Qy       2337 GAGCACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACC 2396

Db        901 GAGCACCACCTCGCCAATGTCAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACC 960

Qy       2397 CTGCAAACGAGGGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGACCCACTCCAACCG 2456

Db        961 CTGCAAACGAGGCTGGGAAGCAAGCTGCTCATCACTGACAGACAGGACCCACTCCACCCG 1020

Qy       2457 ACGGAGACCAGGTTTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCTCCAAT 2516

Db       1021 ACGGAGACCAGGTGTGTTGATGGAAGAGACACTCTGGAGTGGGACATCTGTGCCTCCAAT 1080

Qy       2517 GGGATCATTCATGTCATTTCCAGGCCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGACC 2576

Db       1081 GGGATCACACATGTCATTTCCAGGYCTTTAAAAGCACCCCCTGCCCCCGTGACCTTGNCC 1140

Qy       2577 CACACTGGCTTGGGAGCAGGGATCTTCTTTGCCATCATCCTGGTGACTGGGGCTGTTGCC 2636

Db       1141 CACACTGGNTTGGGAGNAGGGATCTTCTNTGNCATCATCCTGGTGACTGGGGCTGTTGCC 1200

Qy       2637 TTGGCTGCTTACTCCTACTTTCGGATAAACCGGAGAACAATCGGCTTCCAGCATTTTGA 2695

Db       1201 TTGGCTGCTTACTCCTACTTTCGGATAAACCGGAAAACAATCGGCTTCCANCATTTTGA 1259
```

Search completed: June 21, 2004, 10:47:18
Job time : 1193 secs